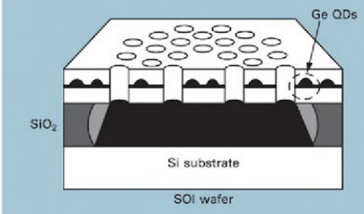
WOODHEAD PUBLISHING IN MATERIALS


Ge QDs
SiO$_2$
Si substrate
SOI wafer

Silicon–germanium (SiGe) nanostructures

Production, properties and applications in electronics

Edited by Yasuhiro Shiraki and Noritaka Usami

WP
WOODHEAD PUBLISHING

VISIT…

LANZAROTE
Caliente.COM

Silicon–germanium (SiGe) nanostructures

Related titles:

High temperature superconductors
(ISBN 978-1-84569-578-1)
High temperature superconductors reviews growth techniques, properties and applications of key high temperature superconducting films. The first part provides an overview of high temperature superconductor properties such as optical conductivity and transport properties. Part II reviews the growth and properties of particular types of superconducting film, whilst Part III describes how they can be applied in practice in various areas of electronics.

Optical switches: materials and design
(ISBN 978-1-84569-579-8)
Optical components are increasingly important in communication networks and significant technological advances have been made in recent years. *Optical switches: materials and design* provides a comprehensive review of this interesting and commercially significant field of research. Written by an international team of expert contributors, the major optical switch technologies are discussed with particular emphasis on the materials, design and associated issues. In addition to introductory material and perspective on research results, topics reviewed include acousto-optical switches, thermo-optical switches, magneto-optical switches, MEMs-based optical switches, liquid crystal optical switches and photonic crystal optical switches.

Advanced adhesives in electronics
(ISBN 978-1-84569-576-7)
Adhesives are widely used in the manufacture of electronic devices to act as passive and active components. Recently there has been considerable interest in the use of conductive adhesives. This book reviews key types of conductive adhesive, processing methods, properties and the way they can be modelled as well as potential applications.

Details of these and other Woodhead Publishing materials books can be obtained by:

- visiting our web site at www.woodheadpublishing.com
- contacting Customer Services (e-mail: sales@woodheadpublishing.com; fax: +44 (0) 1223 832819; tel.: +44 (0) 1223 499140 ext. 130; address: Woodhead Publishing Limited, 80 High Street, Sawston, Cambridge CB22 3HJ, UK)

If you would like to receive information on forthcoming titles, please send your address details to: Francis Dodds (address, tel. and fax as above; e-mail: francis.dodds@woodheadpublishing.com). Please confirm which subject areas you are interested in.

Silicon–germanium (SiGe) nanostructures

Production, properties and applications in electronics

Edited by

Yasuhiro Shiraki and Noritaka Usami

WOODHEAD PUBLISHING

Oxford Cambridge Philadelphia New Delhi

Published by Woodhead Publishing Limited,
80 High Street, Sawston, Cambridge CB22 3HJ, UK
www.woodheadpublishing.com

Woodhead Publishing, 1518 Walnut Street, Suite 1100, Philadelphia, PA 19102–3406, USA

Woodhead Publishing India Private Limited, G-2, Vardaan House, 7/28 Ansari Road,
Daryaganj, New Delhi – 110002, India
www.woodheadpublishingindia.com

First published 2011, Woodhead Publishing Limited
© Woodhead Publishing Limited, 2011
The authors have asserted their moral rights.

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. Reasonable efforts have been made to publish reliable data and information, but the authors and the publisher cannot assume responsibility for the validity of all materials. Neither the authors nor the publisher, nor anyone else associated with this publication, shall be liable for any loss, damage or liability directly or indirectly caused or alleged to be caused by this book.

Neither this book nor any part may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, microfilming and recording, or by any information storage or retrieval system, without permission in writing from Woodhead Publishing Limited.

The consent of Woodhead Publishing Limited does not extend to copying for general distribution, for promotion, for creating new works, or for resale. Specific permission must be obtained in writing from Woodhead Publishing Limited for such copying.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation, without intent to infringe.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

ISBN 978-1-84569-689-4 (print)
ISBN 978-0-85709-142-0 (online)

The publisher's policy is to use permanent paper from mills that operate a sustainable forestry policy, and which has been manufactured from pulp which is processed using acid-free and elemental chlorine-free practices. Furthermore, the publisher ensures that the text paper and cover board used have met acceptable environmental accreditation standards.

Typeset by Replika Press Pvt Ltd, India
Printed by TJI Digital, Padstow, Cornwall, UK

Contents

Contributor contact details *xiii*

Preface *xix*

Part I Introduction

1 Structural properties of silicon–germanium (SiGe) nanostructures 3

E. Kasper and H.-J. Herzog, University of Stuttgart, Germany

1.1 Introduction 3
1.2 Crystal structure 4
1.3 Lattice parameters 4
1.4 Phase diagram 7
1.5 Critical thickness 8
1.6 Structural characterization by X-ray diffraction 17
1.7 Future trends 23
1.8 Acknowledgement 23
1.9 References 24

2 Electronic band structures of silicon–germanium (SiGe) alloys 26

N. Mori, Osaka University, Japan

2.1 Band structures 26
2.2 Strain effects 29
2.3 Effective mass 38
2.4 Conclusion 41
2.5 References 41

Part II Formation of nanostructures

3 Understanding crystal growth mechanisms in silicon–germanium (SiGe) nanostructures 45
M. Suemitsu, Tohoku University, Japan and S. N. Filimonov, Tomsk State University, Russia

3.1 Introduction 45
3.2 Thermodynamics of crystal growth 46
3.3 Fundamental growth processes 50
3.4 Kinetics of epitaxial growth 60
3.5 Heteroepitaxy 64
3.6 References 68

4 Types of silicon–germanium (SiGe) bulk crystal growth methods and their applications 72
N. Usami, Tohoku University, Japan

4.1 Introduction 72
4.2 Growth methods 74
4.3 Application of silicon–germanium (SiGe) bulk crystal to heteroepitaxy 79
4.4 Conclusion 79
4.5 References 80

5 Silicon–germanium (SiGe) crystal growth using molecular beam epitaxy 83
A. Sakai, Osaka University, Japan

5.1 Introduction 83
5.2 Techniques 84
5.3 Nanostructure formation by molecular beam epitaxy (MBE) 87
5.4 Future trends 113
5.5 References 114

6 Silicon–germanium (SiGe) crystal growth using chemical vapor deposition 117
B. Tillack, IHP and Technische Universität Berlin, Germany and J. Murota, Tohoku University, Japan

6.1 Introduction 117
6.2 Epitaxial growth techniques – chemical vapor deposition (CVD) (ultra high vacuum CVD (UHVCVD), low pressure CVD (LPCVD), atmospheric pressure CVD (APCVD), plasma enhanced CVD (PECVD)) 119

6.3 Silicon–germanium (SiGe) heteroepitaxy by chemical vapor deposition (CVD) 122
6.4 Doping of silicon–germanium (SiGe) 128
6.5 Conclusion and future trends 139
6.6 References 140

7 Strain engineering of silicon–germanium (SiGe) virtual substrates 147
K. Sawano, Tokyo City University, Japan

7.1 Introduction 147
7.2 Compositionally graded buffer 148
7.3 Low-temperature buffer 158
7.4 Ion-implantation buffer 161
7.5 Other methods and future trends 165
7.6 Sources of further information 167
7.7 References 167

8 Formation of silicon–germanium on insulator (SGOI) substrates 171
N. Sugiyama and T. Tezuka, Toshiba Corporation, Japan

8.1 Introduction: demand for virtual substrate and (Si)Ge on insulator (SGOI) 171
8.2 Formation of (Si)Ge on insulator (SGOI) by the Ge condensation method 172
8.3 Extension toward Ge on insulator 186
8.4 Conclusion 187
8.5 Acknowledgment 188
8.6 References 188

9 Miscellaneous methods and materials for silicon–germanium (SiGe) based heterostructures 190
M. Miyao, Kyushu University, Japan

9.1 Introduction 190
9.2 Oriented growth of silicon–germanium (SiGe) on insulating films for thin film transistors and 3-D stacked devices 191
9.3 Heteroepitaxial growth of ferromagnetic Heusler alloys for silicon–germanium (SiGe)-based spintronic devices 201
9.4 Conclusion 207
9.5 References 208

10 Modeling the evolution of germanium islands on silicon(001) thin films 211
L. Miglio and F. Montalenti, University of Milano-Bicocca, Italy

10.1 A few considerations on epitaxial growth modeling 211
10.2 Introduction to Stranski–Krastanow (SK) heteroepitaxy 213
10.3 Onset of Stranski–Krastanow (SK) heteroepitaxy 219
10.4 Beyond the Stranski–Krastanow (SK) onset: Si–Ge intermixing 232
10.5 Beyond the Stranski–Krastanow (SK) onset: vertical and horizontal ordering for applications 235
10.6 Future trends: ordering Ge islands on pit-patterned Si(001) 241
10.7 References 243

11 Strain engineering of silicon–germanium (SiGe) micro- and nanostructures 247
F. Pezzoli, C. Deneke and O. G. Schmidt, IFW Dresden, Germany

11.1 Introduction 247
11.2 Growth insights 250
11.3 Island engineering 260
11.4 Rolled-up nanotechnology 268
11.5 Potential applications 272
11.6 Sources of further information and advice 279
11.7 Acknowledgments 282
11.8 References 282

Part III Material properties of SiGe nanostructures

12 Self-diffusion and dopant diffusion in germanium (Ge) and silicon–germanium (SiGe) alloys 299
M. Uematsu, Keio University, Japan

12.1 Introduction 299
12.2 Diffusion mechanism 300
12.3 Self-diffusion in germanium (Ge) 302
12.4 Self-diffusion in silicon–germanium (SiGe) alloys 306
12.5 Silicon–germanium (Si–Ge) interdiffusion 310
12.6 Dopant diffusion in germanium (Ge) 316
12.7 Dopant diffusion in silicon–germanium (SiGe) alloys 323
12.8 Dopant segregation 329
12.9 Conclusion and future trends 330
12.10 Sources of further information and advice 331

12.11 References 332

13 Dislocations and other strain-induced defects in silicon–germanium (SiGe) nanostructures 338
M. L. Lee, Yale University, USA

13.1 Introduction and background 338
13.2 Historical overview 341
13.3 Application of the Thompson tetrahedron to extended defects in silicon–germanium (SiGe) 346
13.4 Current topics 349
13.5 Future trends 354
13.6 Acknowledgments 356
13.7 References 357

14 Transport properties of silicon/silicon–germanium (Si/SiGe) nanostructures at low temperatures 361
A. Gold, Université Paul Sabatier, France

14.1 Introduction 361
14.2 Model, disorder and transport theory 363
14.3 Transport in quantum wells 372
14.4 Transport in heterostructures 379
14.5 Comparison with experimental results 383
14.6 Discussion and future trends 389
14.7 Conclusions 392
14.8 Acknowledgements 393
14.9 References 393

15 Transport properties of silicon–germanium (SiGe) nanostructures and applications in devices 399
F. Schäffler, Johannes Kepler Universität, Austria

15.1 Introduction 399
15.2 Basic transport properties of strained silicon–germanium (SiGe) heterostructures 401
15.3 Strain engineering 403
15.4 Low-dimensional transport 407
15.5 Carrier transport in silicon/silicon–germanium (Si/SiGe) devices 424
15.6 Future trends 427
15.7 References 428

16 Microcavities and quantum cascade laser structures based on silicon–germanium (SiGe) nanostructures 433
J. Xia and Y. Shiraki, Tokyo City University, Japan, and J. Yu, Chinese Academy of Sciences, P. R. China

16.1 Introduction 433
16.2 Germanium (Ge) dots microcavity photonic devices 434
16.3 Silicon–germanium (SiGe) quantum cascade laser (QCL) structures 445
16.4 Conclusions 451
16.5 References 452

17 Silicide and germanide technology for interconnections in ultra-large-scale integrated (ULSI) applications 456
S. Zaima and O. Nakatsuka, Nagoya University, Japan

17.1 Introduction 456
17.2 Formation of silicide and germanosilicide thin films 457
17.3 Crystalline properties of silicides 459
17.4 Electrical properties 462
17.5 References 466

Part IV Devices using silicon, germanium and silicon–germanium (Si, Ge and SiGe) alloys

18 Silicon–germanium (SiGe) heterojunction bipolar transistor (HBT) and bipolar complementary metal oxide semiconductor (BiCMOS) technologies 473
K. Washio, Hitachi Ltd, Japan

18.1 Introduction 473
18.2 Epitaxial growth 474
18.3 Silicon–germanium (SiGe) heterojunction bipolar transistor (HBT) 479
18.4 Silicon–germanium (SiGe) bipolar complementary metal oxide semiconductors (BiCMOS) 486
18.5 Applications in integrated circuit (IC) and large-scale integration (LSI) 490
18.6 Conclusion 493
18.7 References 494

19 Silicon–germanium (SiGe)-based field effect transistors (FET) and complementary metal oxide semiconductor (CMOS) technologies 499
S. Takagi, The University of Tokyo, Japan

19.1 Introduction 499
19.2 Silicon–germanium (SiGe) channel metal oxide semiconductor field effect transistors (MOSFETs) 506
19.3 Conclusion 524
19.4 References 524

20 High electron mobility germanium (Ge) metal oxide semiconductor field effect transistors (MOSFETs) 528
A. Toriumi, The University of Tokyo, Japan

20.1 Introduction 528
20.2 Gate stack formation 529
20.3 Metal oxide semiconductor field effect transistor (MOSFET) fabrication and electron inversion layer mobility 536
20.4 Germanium (Ge)/metal Schottky interface and metal source/drain metal oxide semiconductor field effect transistors (MOSFETs) 544
20.5 Conclusion and future trends 548
20.6 Acknowledgments 549
20.7 References 549

21 Silicon (Si) and germanium (Ge) in optical devices 551
K. Ohashi, NEC Corporation, Japan

21.1 Background 551
21.2 Optical waveguides 552
21.3 Modulators 556
21.4 Photodetectors and photovoltaics 559
21.5 Light sources 565
21.6 Future trends 566
21.7 Sources of further information and advice 567
21.8 References 567

22 Spintronics of nanostructured manganese germanium (MnGe) dilute magnetic semiconductor 575
K. L. Wang and F. Xiu, University of California, Los Angeles, USA and A. P. Jacob, Intel Corporation, USA

22.1 Introduction 575

22.2 Theories of ferromagnetism in group IV dilute magnetic semiconductor (DMS) 578
22.3 Growth and characterizations of group IV dilute magnetic semiconductor (DMS) and nanostructures 581
22.4 Electric field-controlled ferromagnetism 596
22.5 Conclusion and future trends 601
22.6 References 602

Index 610

Contributor contact details

(* = main contact)

Editors

Y. Shiraki
Research Center for Si Nanoscience
Advanced Research Laboratories
Tokyo City University
8-15-1 Todoroki, Setagaya-ku
Tokyo 158-0082
Japan

E-mail: yshiraki@tcu.ac.jp

N. Usami
Institute for Materials Research
Tohoku University
2-1-1 Katahira, Aoba-ku
Sendai 980-8577
Japan

E-mail: usa@imr.tohoku.ac.jp

Chapter 1

E. Kasper
University of Stuttgart
Institute for Semiconductor Engineering
Pfaffenwaldring 47
70569 Stuttgart
Germany

E-mail: kasper@iht.uni-stuttgart.de

Chapter 2

N. Mori
Division of Electrical, Electronic and Information Engineering
Graduate School of Engineering
Osaka University
Suita City
Osaka 565-0871
Japan

E-mail: nobuya.mori@eei.eng.osaka-u.ac.jp

Chapter 3

M. Suemitsu*
Research Institute of Electrical Communication
Tohoku University
2-1-1 Katahira, Aoba-ku
Sendai 980-8577
Japan

E-mail: suemitsu@riec.tohoku.ac.jp

S.N. Filimonov
Department of Physics
Tomsk State University
36 Lenina Avenue
Tomsk 634050
Russia

E-mail: filimon@phys.tsu.ru

Chapter 4

N. Usami
Institute for Materials Research
Tohoku University
2-1-1 Katahira, Aoba-ku
Sendai 980-8577
Japan

E-mail: usa@imr.tohoku.ac.jp

Chapter 5

A. Sakai
Graduate School of Engineering Science
Osaka University
1-3 Machikaneyama-cho
Toyonaka
Osaka 560-8531
Japan

E-mail: sakai@ee.es.osaka-u.ac.jp

Chapter 6

B. Tillack*
IHP
Im Technologiepark 25
15236 Frankfurt (Oder)
Germany

E-mail: tillack@ihp-microelectronics.com

Technische Universität Berlin
HFT4
Einsteinufer 25
10587 Berlin
Germany

J. Murota
Research Institute of Electrical Communication
Tohoku University
2-1-1 Katahira, Aoba-ku
Sendai 980-8577
Japan

Chapter 7

K. Sawano
Tokyo City University
8-15-1 Todoroki, Setagaya-ku
Tokyo
Japan

E-mail: sawano@tcu.ac.jp

Chapter 8

N. Sugiyama* and T. Tezuka
Corporate R&D Center
Toshiba Corporation
Komukai Toshiba-cho 1
Saiwai-ku, Kawasaki
Japan

E-mail: naoharu.sugiyama@toshiba.co.jp

Chapter 9

M. Miyao
Department of Electronics
Kyushu University
744-Motooka, Nishi-ku
Fukuoka 819-0395
Japan

E-mail: miyao@ed.kyushu-u.ac.jp

Chapter 10

L. Miglio and F. Montalenti*
L-NESS and Materials Science Department
University of Milano-Bicocca
Via Cozzi 53
20125 Milan
Italy

E-mail: leo.miglio@mater.unimib.it
francesco.montalenti@unimib.it

Chapter 11

F. Pezzoli*
L-NESS and Materials Science Department
University of Milano-Bicocca
Via Cozzi 53
20125 Milan
Italy

E-mail: fabio.pezzoli@unimib.it

Institute for Integrative Nanosciences
IFW Dresden
Helmholtzstraße 20
01069 Dresden
Germany

C. Deneke and O. G. Schmidt
Institute for Integrative Nanosciences
IFW Dresden
Helmholtzstraße 20
01069 Dresden
Germany

E-mail: c.deneke@ifw-dresden.de
o.schmidt@ifw-dresden.de

Chapter 12

M. Uematsu
School of Fundamental Science and Technology
Keio University
3-14-1 Hiyoshi
Yokohama 223-8522
Japan

E-mail: uematsu@a3.keio.jp

Chapter 13

M. L. Lee
Department of Electrical Engineering
Yale University
PO Box 208284
New Haven, CT 06520-8284
USA

E-mail: minjoo.lee@yale.edu

Chapter 14

A. Gold
UFR-PCA
Université Paul Sabatier
118 Route de Narbonne
31062 Toulouse
France

Centre d'Elaboration de Materiaux et d'Etudes Structurales
29 Rue Jeanne Marvig
31055 Toulouse
France

E-mail: gold@cemes.fr

Chapter 15

F. Schäffler
Institut für Halbleiter- und Festkörperphysik
Johannes Kepler Universität
Linz
Austria

E-mail: friedrich.schaffler@jku.at

Chapter 16

J. Xia* and Y. Shiraki
Research Center for Si Nanoscience
Advanced Research Laboratories
Tokyo City University
8-15-1 Todoroki, Setagaya-ku
Tokyo 158-0082
Japan

E-mail: jinsongxia@gmail.com

J. Yu
Institute of Semiconductors
Chinese Academy of Sciences
P. O. Box 912
Beijing 100083
P. R. China

E-mail: jzyu@red.semi.ac.cn

Chapter 17

S. Zaima and O. Nakatsuka*
Department of Crystalline Materials Science
Graduate School of Engineering
Nagoya University
Furo-cho, Chikusa-ku
Nagoya 464-8603
Japan

E-mail: zaima@alice.xtal.nagoya-u.ac.jp
nakatuka@alice.xtal.nagoya-u.ac.jp

Chapter 18

K. Washio
Central Research Laboratory
Hitachi Ltd
1-280 Higashi-Koigakubo
Kokubunji
Tokyo 185-8601
Japan

E-mail: katsuyoshi.washio.sq@hitachi.com

Chapter 19

S. Takagi
Department of Electrical Engineering and Information Systems
School of Engineering
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo 113-8656
Japan

E-mail: takagi@ee.t.u-tokyo.ac.jp

Chapter 20

A. Toriumi
Department of Materials Engineering
Graduate School of Engineering
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo 113-8656
Japan

E-mail: toriumi@material.t.u-tokyo.ac.jp

Chapter 21

K. Ohashi
Green Innovation Research Laboratories
NEC Corporation
34 Miyukigaoka
Tsukuba
Ibaraki 305-8501
Japan

E-mail: k-ohashi@cb.jp.nec.com

Chapter 22

F. Xiu and K. L. Wang*
Device Research Laboratory
Department of Electrical Engineering
University of California,
Los Angeles, CA 90095
USA

E-mail: wang@ee.ucla.edu

A. P. Jacob
Intel Corporation
Santa Clara, CA 95054
USA

Preface

In 1975, epitaxial growth of SiGe/Si heterostructures was demonstrated by Erich Kasper and his co-workers using vapour deposition under ultra-high vacuum which is now well known as molecular beam epitaxy (MBE). Although there were quite a few people who were interested in this work at that time, some research groups including the editor's (Y. S.) started the research on silicon MBE in Europe, the United States and Japan. Thanks to the development of crystal growth techniques, particularly MBE, the research field of SiGe heterostructures has expanded substantially not only from the viewpoint of device applications but also from the point of materials science since then. This is because SiGe heterostructures have a high potential to improve the state-of-art Si devices particularly very large scale integrated circuits (VLSIs) and add new functions such as optics. Actually, strained Si structures are now implemented in the sophisticated VLSIs already. They also provide a new scientific field of materials properties and crystal growth relating to the lattice mismatch between Si and Ge which causes strain to modify band structures and growth modes.

Strain-induced band modification brings increase of mobility of both electrons and holes, and the advantages of SiGe have been proved by the commercial spread of hetero-bipolar-transistors (HBTs) as well as the superior performance of strained Si CMOS. Quantum wells based on SiGe/Si heterostructures or quantum dots grown by the Stranski-Krastanov mode make it possible to confine carriers, leading to efficient light emission even with indirect band-gap materials. Optical devices based on highly efficient light emission will provide a new possibility to enhance the performances and functions of Si VLSIs by realizing optical interconnection and optoelectronic integrated circuits (OEICs). Furthermore, SiGe spintronics have attracted enormous attention due to unique magneto-electro-optical properties. As illustrated by these examples, SiGe opens up the prospects of novel and enhanced device performances especially when structural control in nanometre scale is achieved.

This book is intended to provide a comprehensive overview on current understanding on materials science, technology and applications of SiGe nanostructures. The target of this book is the research community in both industry and academia.

The introductory Part I covers the structural properties of SiGe nanostructures, with a further chapter discussing the electronic band structures of SiGe alloys. Part II concentrates on the formation of SiGe nanostructures, with chapters on different methods of crystal growth such as molecular beam epitaxy and chemical vapour deposition. This part also includes chapters covering strain engineering and modelling. Part III covers the material properties of SiGe nanostructures, including chapters on such topics as strain-induced defects, transport properties and microcavities, and quantum cascade laser structures. In Part IV, devices utilizing SiGe alloys are discussed. Chapters cover ultra large scale integrated applications, MOSFETs and the use of SiGe in different types of transistors and optical devices, and spintronics of nanostructured dilute magnetic materials and silicide/silicon heterostructures.

In the process of editing this book, we have received generous assistance from our colleagues. Especially, we are indebted to Erich Kasper for his advice on the organization of the chapters. We extend our sincere thanks to Nell Holden, Benjamin Hilliam, Adam Hooper, Cliff Elwell at Woodhead Publishing Limited for their patience and continuous support.

Yasuhiro Shiraki
Noritaka Usami
Tokyo and Sendai

1

Structural properties of silicon–germanium (SiGe) nanostructures

E. KASPER and H.-J. HERZOG, University of Stuttgart, Germany

Abstract: The heterostructure SiGe/Si has contributed to a large extent to an understanding of lattice mismatched heterostructures and this understanding has led to rapidly increasing exploitation of SiGe in modern microelectronics. In this chapter crystallographic data of silicon–germanium alloys such as crystal structure and lattice parameters and the phase diagram are reviewed. The basic concepts of equilibrium strain and strain relaxation by misfit dislocations are described in the section on critical thickness. The lattice mismatch either causes strain or results in generation of misfit dislocations at the interface. X-ray diffraction is unrivalled as a tool to analyze heteroepitaxial layers.

Key words: heterostructure, silicon–germanium alloy, lattice structure, phase diagram, lattice mismatch, strain, misfit dislocations, critical thickness, X-ray diffraction.

1.1 Introduction

A reliable set of structural data is essential for many investigations of both epilayers, and bulk material. Even the structural assessment and analysis of SiGe epilayers, which are presently of enormous interest for novel and high-performance device applications, require the knowledge of sufficiently precise material data. In this chapter some crystallographic data of silicon–germanium alloys, such as crystal structure and lattice parameters and the phase diagram, are reviewed. For a more complete collection of data the reader is referred to standard volumes on physical properties of semiconductors, e.g. the Landolt-Börnstein Series [1, 2], or specific data reviews, e.g. the EMIS Datareview Series [3, 4].

The technically important structure SiGe/Si serves also as a model system for lattice mismatched heterostructures because chemical effects are less pronounced than in systems with elements from different columns of the periodic table. The basic concepts of equilibrium strain and strain relaxation by misfit dislocations are described in the section on critical thickness. The conceptual structure of this section follows that given in a lecture on ‘Semiconductor Technology’ by one of the authors (Erich Kasper).

1.2 Crystal structure

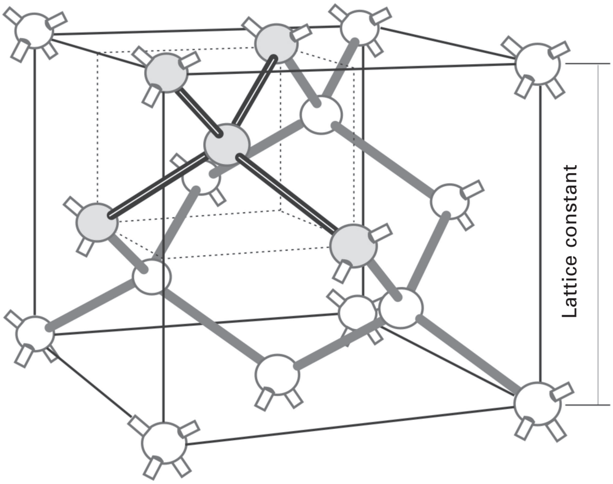
Silicon and germanium, which both crystallize in the diamond lattice, are completely miscible, forming $Si_{1-x}Ge_x$ solid solutions with x ranging from 0 to 1. The space lattice of diamond consists of two face-centred-cubic (fcc) lattices which are displaced a quarter of the space diagonal. A perspective drawing of the unit cell is depicted in Fig. 1.1. The space group of the diamond structure is 0h–Fd3m. The cubic unit cell contains eight atoms that occupy the following positions:

$$\begin{array}{cccccccccccc} 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array}$$

The fractions denote the height above the base in units of the cube edge. In this structure each atom is bonded to four nearest neighbours with a distance of $\sqrt{3} \times a/4$ arranged at the corners of a regular tetrahedron and to 12 next-nearest neighbours. Four tetrahedra form the non-primitive unit cell. The diamond structure is the result of the covalent bonding between the atoms represented by the rods in Fig. 1.1. The diamond lattice is not very compact. Only 34% of the available space is filled with hard spheres.

1.3 Lattice parameters

To date, the most precise and comprehensive determination of bulk lattice parameters (and densities) across the whole $Si_{1-x}Ge_x$ system has been carried



1.1 Diamond crystal structure. Each atom is tetrahedrally bonded to its four nearest neighbours as displayed by the rods.

out by Dismukes *et al.* [5], including measurement of the variation of lattice parameters with temperature up to 800°C for some alloys. In Table 1.1 the lattice parameters of $Si_{1-x}Ge_x$ alloys at 25°C are listed for composition intervals of 5 at% Ge. The data reveal a small deviation from Vegard's law, which means that the SiGe alloy parameters are determined by a linear interpolation of the parameters of the end-point elements Si and Ge.

The departure Δ from Vegard's law, defined by

$$\Delta = a(Si_{1-x}Ge_x) - \{a(Si) + [a(Ge) - a(Si)]x\} \qquad 1.1$$

with x = atomic fraction of Ge, is also listed in Table 1.1. Δ is negative throughout the system with a broad maximum in the middle of the system. This slight deviation from Vegard's law has been confirmed in a recent study on SiGe epitaxial layers on Si(100) substrates analysed by X-ray diffraction and Rutherford backscattering [6]. The experimentally determined deviation from Vegard has also been found theoretically by means of Monte Carlo simulations on $Si_{1-x}Ge_x$ alloys [7, 8].

By using the values given in [5] for $x = 0\%$, $x = 25\%$ and $x = 100\%$, a parabolic relation for the $Si_{1-x}Ge_x$ lattice parameter as a function of the Ge fraction x,

Table 1.1 Lattice parameter a of $Si_{1-x}Ge_x$ alloys for x from 0 to 100 at% in 5% steps after [5]. The right column gives the deviation from Vegard's law Δ

X (at% Ge)	$a(Si_{1-x}Ge_x)$ (nm)	Δ
0	0.54310	…
5	0.54410	– 0.00004
10	0.54522	– 0.00014
15	0.54624	– 0.00026
20	0.54722	– 0.00041
25	0.54825	– 0.00051
30	0.54928	– 0.00062
35	0.55038	– 0.00065
40	0.55149	– 0.00067
45	0.55261	– 0.00068
50	0.55373	– 0.00069
55	0.55492	– 0.00063
60	0.55609	– 0.00060
65	0.55727	– 0.00055
70	0.55842	– 0.00053
75	0.55960	– 0.00048
80	0.56085	– 0.00027
85	0.56206	– 0.00023
90	0.56325	– 0.00019
100	0.56575	…

$$a(\mathrm{Si}_{1-x}\mathrm{Ge}_x) = 0.002733x^2\ (\mathrm{nm}) + 0.01992x\ (\mathrm{nm}) + 0.5431\ (\mathrm{nm}) \tag{1.2}$$

can be derived which approaches the experimental data with a maximum deviation of about 0.0001 nm.

An even better fit is possible by a cubic approach as, for example, by the relation

$$a(\mathrm{Si}_{1-x}\mathrm{Ge}_x) = a(\mathrm{Si})x + a(\mathrm{Ge})(1-x) - 0.00436x^3 + 0.03265x^2 - 0.02829x \tag{1.3}$$

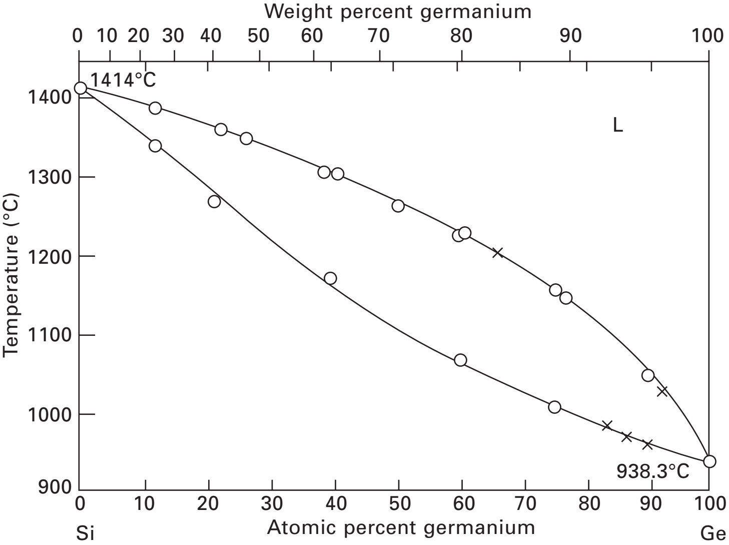
which is implemented in the Philips simulation software for high-resolution X-ray diffraction [9]. The lattice parameter of Si of $a(\mathrm{Si}) = 0.5431$ nm is confirmed by high-precision measurements on pure single crystal Si [10]. The published room-temperature data on the lattice parameter of undoped Ge single crystal range from $a(\mathrm{Ge}) = 0.56573$ nm [11] to 0.56579 nm [12].

In this context, it is to be noted that although SiGe is a well-known alloy system crystallizing in the rather simple diamond lattice, and although Si/SiGe heteroepitaxy structures presently find increasing application due to their promising potential for high-performance devices, the local atomic arrangement such as bond lengths and bond angles is still an object of controversial discussion. There are two extreme concepts for the correlation between bond configuration and alloy composition, termed Bragg/Pauling's and Vegard's limit. According to Bragg [13] and then to Pauling and Huggins [14] the bond lengths in alloys are the sum of the atomic radii of their constituents and remain unchanged as composition varies. Consequently, the bond angles must change to accommodate the presence of different atoms. Vegard's limit [15], on the other hand, means that the bond angles are fixed at the tetrahedral angle and the bond lengths change linearly with the composition. In most experimental studies done by X-ray diffraction (XRD) and in particular by (extended) X-ray absorption fine-structure ((E)XAFS) analysis it has been found that SiGe alloys are mostly, but not completely, Pauling-type in nature, which is in disagreement with nearly all published theoretical results. A comprehensive overview on this topic is given in a recent paper by Aubry *et al.* [16].

Concerning the composition dependency of further structural parameters such as elastic moduli covering the whole composition range, the number of publications in the literature is rather limited. The elastic constants, for example, are required if the lattice parameters of heteroepitaxial $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ films have to be corrected for elastic strain. Mendik *et al.* [17] investigated $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ alloys and found experimental values of the elastic constants C_{ij} which are larger than those calculated from a linear combination of the C_{ij} values of the pure constituents (Table 1.2). From Raman measurements on $\mathrm{Si/Si}_{0.52}\mathrm{Ge}_{0.48}$ strained layer superlattices on Si, Zhang *et al.* [18] obtained for the Si layer a smaller sound velocity and density compared with the Si bulk values, and for the SiGe alloy layers a higher sound velocity and

Table 1.2 Elastic stiffness constants C_{ij} of Si and Ge. For SiGe a linear interpolation is recommended

Stiffness constant (GPa)	Si	Ge
C_{11}	165.8	128.5
C_{12}	53.9	48.3
C_{44}	79.6	66.8



1.2 Liquidus–solidus curve of the $Si_{1-x}Ge_x$ system after [21]. The circles and the crosses are taken from [18] and [19], respectively.

higher density than the values deduced by linear interpolation. However, these changes may possibly be due to strain introduced during formation of the superlattice.

1.4 Phase diagram

The phase diagram in Fig. 1.2 was established by elaborate thermal and X-ray analysis [19]. It is still the basis of the solidification curve of the SiGe system and only a few points have been later added to the Ge-rich side of the phase diagram [20]. No phase changes or decomposition were detected by X-ray analysis after annealing homogeneous solid solution crystals for several months at temperatures in the range from 177°C to 925°C. The SiGe system is a typical representative for a system with strong segregation, i.e. for solid solutions in which the solid and liquid phases are separated by a large regime of coexistence. From this it is evident that, for example, the preparation of a homogeneous solid solution from Si and Ge requires considerable effort because during solidification from the molten phase the Si component strongly segregates and thus quasi-decomposition occurs. Under the assumption that Si and Ge form ideal liquid and solid solutions, Thurmond [21] calculated the liquidus and solidus curves and found the latter to coincide with that in Fig. 1.2. The liquidus lies only slightly below the experimental curve. Some structural and thermodynamic data on the SiGe system are presented in [22] and an early review on the bulk growth of SiGe solid solutions is presented in [23]. Recent advances in the bulk growth of SiGe are summarized in Chapter 4 of this book.

1.5 Critical thickness

Heterostructure device concepts promise strong advantages in micro- and optoelectronics due to gainful effects of band line-up and strain on carrier transport. From the material point of view, the main obstacle to be overcome is the large lattice mismatch of silicon-based heterostructures. One of the best of them, silicon germanium (SiGe), is lattice mismatched to silicon by up to 4% depending on its Ge content. Basic investigations on strained layer growth, interface properties, and deviation from equilibrium are done with SiGe/Si heterostructures.

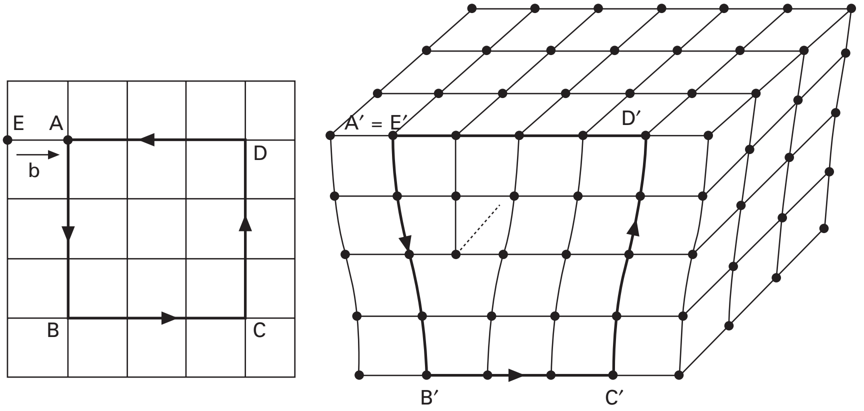
Within the historical development of microelectronics the community had to learn that serious technical obstacles had to be overcome in order to benefit from the heterostructure advantages. The most serious of these obstacles were chemical differences and lattice-mismatch. A model system for lattice-matched heterosystems was found in the III–V material realm with GaAs/GaAlAs nearly perfectly fitting together. But joining III–V semiconductors with silicon (Si) turned out to be difficult because of the chemistry defining each of the materials as dopants for the other group. The main attention was therefore focused on the group IV elements and compounds. In Table 1.3 the cubic (diamond or zincblende lattice cell) group IV compounds are summarized.

The differences in lattice constants are very large with the exception of the completely miscible alloy SiGe. In equilibrium only small amounts of carbon in Si and of tin in Ge are soluble. With the advent of low-temperature growth techniques like molecular beam epitaxy (MBE) or advanced chemical vapour deposition (CVD) methods, the formation of non-equilibrium compounds such as SiGe:C and SiGe:Sn seemed possible. Recently the incorporation of Sn in the Ge lattice has been attracting considerable attention [24–26] because already with 10% Sn a direct (group IV) semiconductor material is predicted and, on top of Ge:Sn tensile strained Ge, is expected to exhibit spectacular carrier transport properties.

In the following part of the chapter, the impact of lattice mismatch and

Table 1.3 Group IV elements and compounds diamond (C), silicon carbide (SiC), silicon (Si), germanium (Ge) and tin (Sn), showing their atomic number (*Z*), atomic weight (*M*), density (ρ), lattice constant a_0 and band gap E_g (eV) at 0 K

Compound	Diamond (C)	SiC	Si	Ge	Sn
Z	6	14; 6	14	32	50
M	12	28; 12	28.1	72.6	118.7
ρ (kg/m^3)	2620	3200	2330	5320	7300
a_0 (pm)	356.7	435.9	543.1	565.8	648.9
E_g (eV)	5.48	2.6	1.17	0.74	0.0

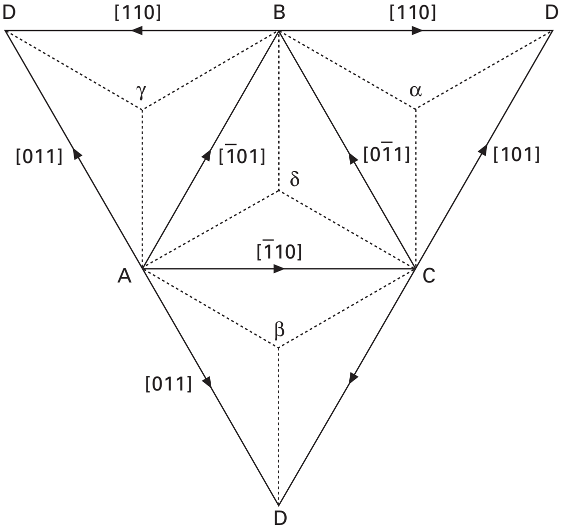


1.3 A positive edge dislocation formed by inserting an extra half-layer of atoms between atomic planes above the slip plane. The Burgers vector of a dislocation is represented as a closure failure EA = b in the ideal lattice when a closed circle was drawn around the dislocation in a real lattice.

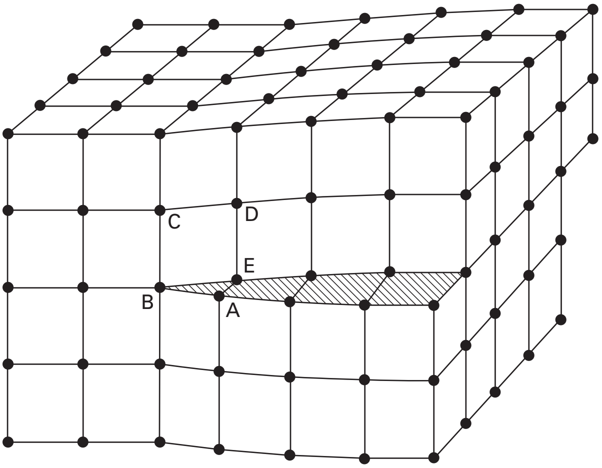
strain on nanometre layers and their electronic structure is evaluated. Surface morphology and misfit dislocation networks play an important role in strain relief processes. The basics of dislocations were introduced in metallurgy where plastic deformation was explained as an atomic phenomenon by dislocation movement. For the reader not so familiar with dislocation concepts they are listed below.

In the diamond cubic lattice the dislocations propagate primarily by glide on {111} planes. The dislocation (Fig. 1.3) is a line defect which is characterized by its local line vector l and its characteristic Burgers vector b, the nature of which is indicated by an atom-by-atom circuit along the path around the dislocation. The dominant Burgers vector in the diamond lattice is along 1/2<110>, and the Burgers vector length b is $a/\sqrt{2}$ (0.38 nm in Si). A dislocation cannot end in the interior of a crystal; it either ends at a surface or is closed as a dislocation loop. Under undisturbed conditions the dislocation follows preferentially the <110> directions. A simple construction due to Thompson [27] greatly assists in the understanding of interactions between dislocations (Fig. 1.4). The tetrahedron ABCD is made up of the four {111} planes. For convenience this tetrahedron is opened, the vertex D having originally been above the plane of the paper. In Fig. 1.3 the Burgers vector is perpendicular to the dislocation line, which is referred to as an edge dislocation. Dislocations with Burgers vector parallel to the dislocation line are known as screw dislocations (Fig. 1.5). Other angles between *l* and *b* exist; these are mixed dislocations, e.g. 60° dislocations of AB dislocation lines with AD Burgers vector in Thompson's reference tetrahedron. Dislocations may easily glide along the {111} slip planes; the slip plane is defined by both vectors – the line direction *l* and Burgers vector *b*. Movement of the dislocation perpendicular to the glide plane is called climb, which needs generation or annihilation of point defects and is therefore much slower. However, there are four {111} planes and so a dislocation can change to another plane under certain circumstances; this is called cross-slip: see, e.g., in Thompson's tetrahedron the dislocation with AD Burgers vector, which may move within the ACD plane and also within the ADB plane (Fig. 1.4).

Despite the severe local distortion around the dislocation, there is still matching of most lattice planes (except the one which causes the displacement *b*). The strain decreases with increasing distance from the dislocation such that displacement of atoms a few atomic spacings away from the dislocation



1.4 Thompson's reference tetrahedron.



1.5 A right-handed screw dislocation formed by displacement of a part of the crystal above the slip-plane parallel to the dislocation line.

may be treated using linear elasticity theory. So the energy per unit length $E_{ds}$of a single dislocation is given by

$$E_{ds} = \frac{\mu b^2}{4\pi}\ln\left(\frac{R}{r_i}\right) \quad 1.4$$

for a screw dislocation and

$$E_{ds} = \frac{\mu b^2}{4\pi(1-\nu)}\ln\left(\frac{R}{r_i}\right) \quad 1.5$$

for an edge dislocation, where μ = shear modulus, ν = Poisson's number, and R and r_i are the outer and inner cut-off radii, respectively.

The energy dependence on the square of the Burgers vector length b causes the Burgers vector to be usually the smallest possible one: 1/2<110>. Smaller vectors like a/6<112> for partial dislocations are indeed possible, but are connected with a plane defect (a stacking fault). In the Thompson tetrahedron (Fig. 1.4) the partial dislocation vectors are given by the dashed lines (e.g. A–δ). The core of the dislocations is often split up in two partials with a stacking fault between them (e.g. Aβ + βD for an AD Burgers vector). This complication of the dislocation core is not considered further in this short review. Since the strain energy is very large, dislocation lines cannot exist in thermal equilibrium in uniform solids despite the increased configurational entropy. In heterostructure couples a dislocation network at the interface may exist in equilibrium because the homogeneous strain in the heterostructure is reduced. The energy balance is the starting point for equilibrium theories, the manifold, slightly differing results from them being caused by the choices of the cut-off radii R and r_i, respectively. The inner cut-off radius is of the order of the Burgers vector length b. The outer radius is given by the geometry of the sample as shown for different geometries by van der Merwe's group. In misfit dislocation theory of thin films on thick substrates, two situations are especially interesting: for low densities of misfit dislocations the outer cut-off is given by the film thickness t and for high densities it is given by the half-distance $p/2$ to the neighbouring dislocation line. From the energy of a dislocation a force F can be deduced; the resolved shear stress on a glide plane and the straightening tendency of curved lines are then transformed into forces. A very popular equilibrium theory of misfit dislocations from Matthews and Blakeslee is based on the force consideration.

The force F_σ of the shear stress is given by

$$F_\sigma = 2\mu(b\cos\lambda)h\varepsilon\frac{1+\nu}{1-\nu} \qquad 1.6$$

which tends to extend the misfit dislocation line, whereas a restoring line tension force F_T acts from the bent arrangement of misfit dislocation and connected threading dislocation:

$$F_T = \mu b^2\frac{(1-\nu\cos^2\theta)}{4\pi(1-\nu)}\ln\left(\frac{h}{r_i}\right) \qquad 1.7$$

This formula accounts for the mixed character of the dislocation (the angle θ between l and b reflects the edge or screw character, $\cos\theta = \frac{1}{2}$ for the common 60° dislocation) and for the effective misfit strain relaxation (the angle λ between the Burgers vector b and line normal in the interface reflects the interface edge component of the dislocation which defines the amount of lattice adjustment, $\cos\lambda = \frac{1}{2}$ for the 60° dislocation). Screw components and edge components perpendicular to the interface contribute to a twist-and-tilt of the layer, respectively. These components usually cancel each other out on the four symmetric {111} slip planes on a (100) oriented sample.

The lattice mismatch either causes strain or results in generation of misfit dislocations at the interface. The thickness t at which strain is partly relaxed by the first misfit dislocations is called the critical thickness t_c. Its value is well defined only at thermal equilibrium whereas at low growth temperatures (e.g. 550°C) the nucleation and glide of dislocations is kinetically impeded, resulting in a rather large metastable region. By lowering the growth temperature below 550°C the critical thickness may be further increased, in a region which is called ultrametastable [6].

The lattice constant of SiGe is slightly larger than that of Si. For a rough estimate the lattice mismatch f of an alloy is linearly interpolated (Vegard's law) between the two parent materials:

$$f = \frac{a_f - a_s}{a_s} = 0.042x \qquad 1.8$$

where a_f and a_s are the lattice constants of film and substrate, respectively, and x is the Ge molar content of the alloy.

For more exact calculations a small parabolic deviation has to be considered. Recent investigations [6] on epitaxial material confirmed the old measurements on bulk crystals [5]. Taking into account the parabolic deviation, the lattice mismatch f of SiGe reads exactly as:

$$f = 0.0417x - 0.005x(1 - x) \qquad 1.9$$

The structural and morphological stability of lattice-mismatched heterostructures is rather complex and covers both extremes, instability as well as perfect stability. For device processing a basic insight into the fundamental mechanisms of mismatch accommodation is necessary. Nature gives several answers to lattice mismatch. In the case of SiGe on Si the answers are strained films, strain relaxation by misfit dislocations, surface undulations, and cracking. Cracking happens only during cooling down of thick Ge films and does not needs to be considered for most devices. To simplify the picture, we first treat the equilibrium answer and then discuss the kinetic limitations imposed by the low growth temperatures and by processing of multilayer device structures.

Up to the critical thickness t_c the layers are strained, which means that every atomic row in the substrate is continuing across the interface. This low-thickness regime is completely stable. The reason is given by the energy balance favouring strain in a small volume instead of creating possible atomic defect structures like misfit dislocations. Misfit dislocations are line defects where atomic rows do not continue across the interface. Misfit dislocations are characterized by the line vector l and the Burgers vector b which defines the inhomogeneous strain field around the dislocation. The glide plane given by the vectors l and b allows easy movement of the dislocation. In diamond lattices the glide plane is the densely packed (111) plane. Dislocations cannot end inside the material; they are either closed or end at surfaces.

Let us first consider a strained layer on top of the substrate. The in-plane strain value ε is directly linked to the mismatch f:

$$\varepsilon = -f\frac{a_s}{a_f} \cong -f \quad \text{(strained layer)} \qquad 1.10$$

The minus sign in the equation results from the given definitions of mismatch f and strain ε. A larger-layer film has a positive mismatch; the resulting compressive strain is referred to as negative. The near-unity correction factor a_s/a_f stems from the different reference lattices. For mismatch definition the substrate lattice is usually used as reference whereas for the film strain the film lattice is the reference. In what follows we take the correction factor as unity.

The lattice expansion by one dislocation is given by the effective Burgers vector $b' = b\cos\lambda$. The Burgers vector is only effective for lattice accommodation with its edge-type component in the interface plane. The angle λ is therefore defined as the angle between the Burgers vector and the dislocation-line normal vector in the interface.

The misfit dislocation density of a planar arrangement is defined by $(1/p)$, the reciprocal of the dislocation spacing p (note that the dislocation density within a volume – e.g. for threading dislocations – is given by dislocation lengths/volume, which reads as dislocation per area, whereas the density of dislocations in a sheet is given by dislocation lengths/area, which reads as dislocation per length). For a completely relaxed film the misfit dislocation spacing p in the interface is given by (b' taking the same sign as f):

$$p = \frac{b'}{f} \tag{1.11}$$

For a partly relaxed film with both dislocations and strain, relation 1.12 holds:

$$f = -\varepsilon + \frac{b'}{f} \tag{1.12}$$

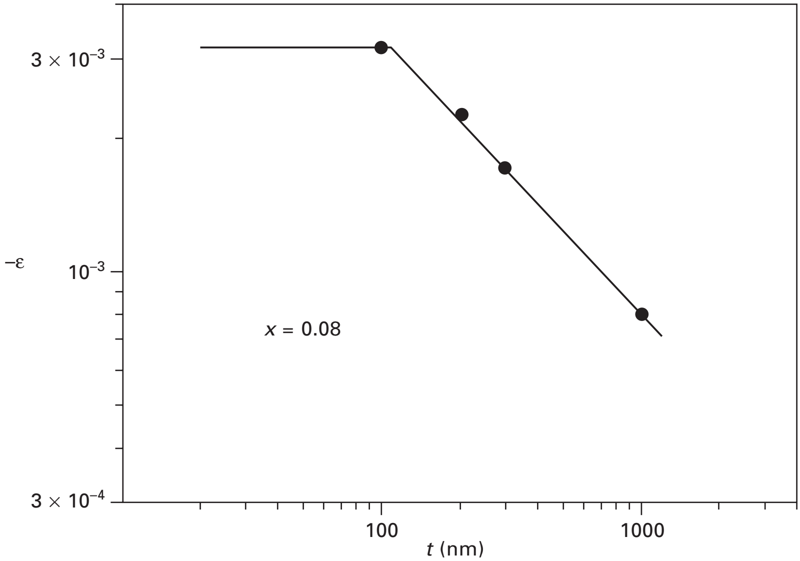
The relaxation degree r describes the relative amount of strain released by the introduction of misfit dislocations:

$$r = 1 + \left(\frac{\varepsilon}{f}\right) = 1 - \left|\frac{\varepsilon}{f}\right| \tag{1.13}$$

The relaxation starts with a film of the critical thickness t_c. With further increase of the layer thickness the strain decreases (Fig. 1.6). Full relaxation ($r \rightarrow 1$, $\varepsilon \rightarrow 0$) is approached only at thicknesses much larger than the critical thickness [28].

Above the critical film thickness (≈0.1 μm), the compressive strain is lowered by misfit dislocations. The equilibrium critical thickness may be calculated rather easily from basic dislocation theory when – the reader should keep this in mind – the film is assumed to be flat. Differences in published numerical values are caused by differing assumptions about the dislocation core and the range of the inhomogeneous strain [4]. The as-calculated equilibrium thickness t_{cm} (the subscript m refers to van der Merwe and Matthews–Blakeslee [29] who pioneered this kind of equilibrium calculation) is rather small, e.g. 6 nm for a 25% SiGe alloy ($f \cong 0.01$):

$$\left(\frac{t_{\mathrm{cm}}}{b}\right) f = 5.78 \times 10^{-2} \ln\left(\frac{t_{\mathrm{cm}}}{b}\right) \tag{1.14}$$

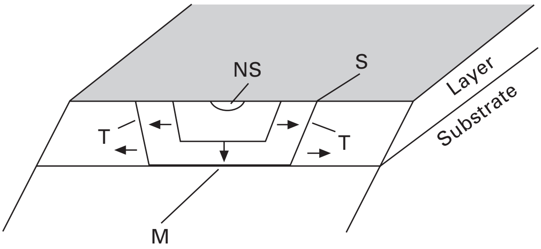


1.6 Compressive film strain vs film thickness *t* of $Si_{0.92}Ge_{0.08}$ layers on Si substrates.

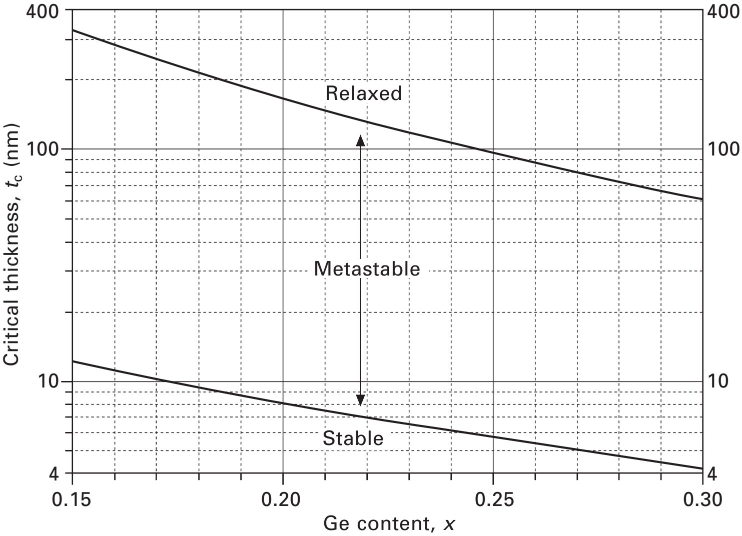
The magnitude *b* of the Burgers vector is 0.384 nm in Si.

The concept of the equilibrium critical thickness in the simple form given here is only meaningful for low lattice mismatches. At higher mismatch values the film strain is reduced by surface corrugations which cause a conversion from the flat van der Merwe growth mode to direct island growth (Volmer–Weber mode) or island growth on a thin wetting layer (Stranski–Krastanov mode), e.g. Ge on Si builds a wetting layer only about 0.5 nm thick on which nucleation of islands takes place. When the islands grow larger then misfit dislocations are generated additionally. The generation of at least the first dislocations is assumed to be from surface nucleation sites (NS) which are either small imperfections or surface steps (Fig. 1.7). Under the strain force, dislocation half-loops move to the interface, creating a straight misfit dislocation segment (M) and two threading dislocation arms (T) connecting the misfit dislocation with the surface. The misfit dislocation network when placed below the active device could be the ultimate near-device getter, but the threading dislocations cause electronic problems with this kind of material.

The growth temperature of SiGe is strongly reduced (typically 450–750°C) compared to Si epitaxy. Surface atom migration and dislocation nucleation are kinetically suppressed. Flat, strained layers (pseudomorphic structure) are obtained under metastable growth conditions with much higher thicknesses and higher mismatch values than under equilibrium. In Fig. 1.8 the critical



1.7 Nucleation of dislocation half-loops (nucleation site NS) and movement to the interface, creating segments M (misfit dislocation) and T (threading dislocation).



1.8 Critical thickness t_c as a function of the Ge content *x*. Stable region (equation 1.7) and metastable region (equation 1.8) are strained. Relaxed region with misfit dislocations.

thickness under equilibrium (lower line) and that obtained at low growth temperatures (MBE, 550°C) are compared.

The metastable critical thickness t_{cb} (the subscript b refers to People and Bean [30] who discovered the large amount of metastability and published the fit relation in 1986) is roughly given by [4]:

$$\left(\frac{t_{cb}}{b}\right)f^2 = \frac{1}{200}\ln\left(\frac{t_{cb}}{b}\right) \qquad 1.15$$

With growth temperatures below 550°C even higher critical thicknesses can be obtained (ultrametastability).

Strained layers just below the equilibrium critical thickness t_{cm} are inherently stable. But also metastable layers turned out to be stable against heat treatments considerably higher than their growth temperature. The main reason is given by the silicon cap on top of the SiGe structure which is grown either for functional purposes (emitter of the heterobipolartransistor) or to facilitate technological steps (oxidation, resist coverage). The equilibrium critical thickness is doubled by the cap but the kinetic limitation is even stronger, as one can easily understand that from a surface nucleation site the dislocation half-loops are not forced to move through the unstrained cap. Processing of strained SiGe up to 850°C is successfully demonstrated.

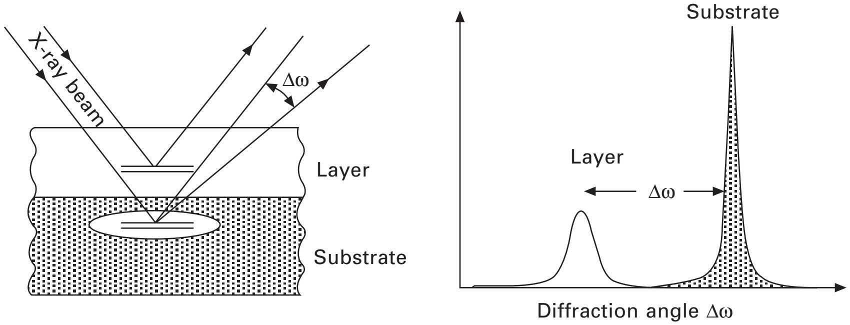
Direct observation of misfit dislocations is possible by X-ray topography and transmission electron microscopy (TEM). Above the critical thickness, dislocations nucleate and move to the interface, leading to an increasing misfit dislocation density with increasing layer thickness. In plane-view TEM micrographs the misfit dislocations are seen as mainly straight lines along both perpendicular <110> directions of the [101] interface [28]. Long misfit dislocation segments and thus a low density of threading dislocations are typical for high growth temperatures and low-mismatched layers.

The slow relaxation r with increasing thickness defines a broad range of misfit and thickness, where both dislocations and strain have to be considered.

1.6 Structural characterization by X-ray diffraction

X-ray diffraction and in particular high resolution X-ray diffraction (HR-XRD) are unrivalled tools to analyse the crystal structure of heteroepitaxial layers. The key feature of HR-XRD is the monochromatic and parallel X-ray beam. The optimal solution to achieve such an X-ray beam with a wavelength λ is a four-crystal monochromator (type duMond–Hart–Bartels) placed between the X-ray source and the sample goniometer. The standard measurement procedure is to record the diffracted intensity as a function of the angular position of the sample and the detector (θ–2θ scans), usually termed the X-ray rocking curve (RC). A difference Δd between the lattice net planes with the Miller indices (hkl) of an epilayer and the substrate leads to two fundamental diffraction lines in the RC as sketched in Fig. 1.9. The relation between Δd and the angular spacing $\Delta\omega$ is given by the well-known Bragg's law, $\lambda = 2d\sin\vartheta$, leading to the derivative form $\Delta d/d = -\cot\vartheta \times \Delta\omega$.

Generally, the X-ray diffraction pattern is the Fourier transform of the carrier density distribution of the probed sample structure. Therefore, high resolved RCs not only enable the determination of lattice parameters but also contain information on thickness and crystal quality of epilayers on substrates. Moreover, as the technique is completely non-destructive it can be applied at different stages of a processed sample.



1.9 X-ray diffraction from a single crystal heterostructure leads to two fundamental diffraction lines in the diffractogram with a distance $\Delta\omega$ of typically a few hundred to a few thousand seconds of arc.

Silicon and germanium are both diamond lattices. If a SiGe layer is grown on a Si(100) substrate (the standard material in Si technology) there is a lattice mismatch $f = (a_L - a_S)/a_S$ between the natural lattice constants a_L and a_S of the layer and the substrate, respectively. The unit cell of a SiGe epilayer on Si can accommodate this mismatch by three modes: (i) fully strained, (ii) partly relaxed, and (iii) fully relaxed. Only in case (iii) is the layer not strained and maintains its natural cubic lattice parameter a_L(SiGe). In cases (i) and (ii) the layer is biaxially in-plane strained, which leads to a tetragonal deformation of the lattice. Its original cubic unit cell a_L splits into lattice parameters parallel (a_T) and perpendicular (c_T) to the interface:

$$a_T = a_L(1 + \varepsilon^{\|}) \tag{1.16}$$

$$c_T = a_L(1 + \varepsilon^{\perp}) \tag{1.17}$$

with $\varepsilon^{\|}$ and $\varepsilon^{\perp}$ being the in-plane and perpendicular strain components, respectively. Assuming the validity of Hooke's law, the strain components are correlated via the relation

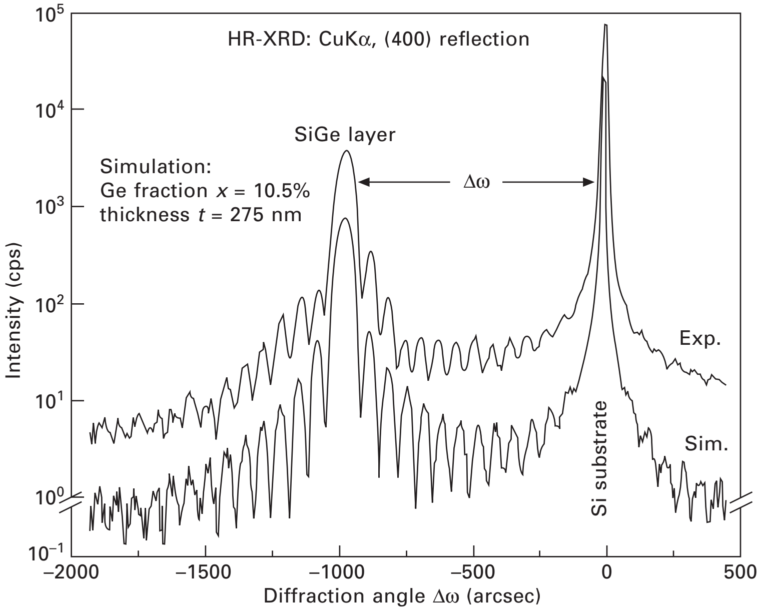
$$\varepsilon^{\perp} = -\frac{2\nu}{1-\nu}\,\varepsilon^{\|} \tag{1.18}$$

where ν is the Poisson ratio, which can be written in terms of the elastic constants c_{11} and c_{12}:

$$\nu = c_{12}/(c_{11} + c_{12}) \tag{1.19}$$

In the case of a SiGe layer fully strained on Si(100) substrate ($\varepsilon^{\|} = |f|$) just one symmetric RC (e.g. diffraction on surface-parallel net planes with Miller indices (400)) is sufficient to analyse the tetragonally deformed unit cell of the layer. The in-plane lattice parameter is equal to that of the substrate, and the perpendicular parameter can be simply obtained from the angular distance $\Delta\theta$ between the substrate and the layer peak according to Bragg's law. With equations 1.16–1.18, the natural lattice parameter of the SiGe alloy layer is calculated, which can then be converted into the corresponding Ge fraction by using the data of Table 1.1. A more state-of-the-art evaluation of the layer parameters consists in the application of a computer program based on dynamic diffraction theory. By means of a fit procedure between a simulated model layer structure and the experimental RC, layer composition and layer thickness are obtained with very high accuracy. An example of such a procedure is given in Fig. 1.10. It shows the experimental HR-XRD plot of a SiGe layer on Si(100) and the corresponding computer-simulated fit. In addition to the fundamental Si substrate and the SiGe layer peak, on both sides of the SiGe line satellites appear. These are so-called thickness fringes which are a precise measure for the layer thickness of the SiGe layer. Moreover, the appearance of thickness oscillations is also a sensitive indicator for good crystal quality, as defect-induced incoherent X-ray scattering rapidly leads to a destruction of the fringes. A very good agreement is obtained with the parameter set of 275 nm SiGe with 10.5% Ge on Si(100).

Several manufacturers that produce HR-XRD machines have simulation software in their product range, such as, e.g., X'Pert Epitaxy from PANalytical [9], JV-HRXRD analysis software from Jordan Valley Semiconductor [31]



1.10 X-ray diffractogram of a SiGe layer on Si(100) (exp.) compared to a computer simulated SiGe on Si structure (sim.) with 10.5% Ge fraction and 275 nm thickness.

and LEPTOS from Bruker AXS [32]. For Fig. 1.10 the program PC-HRS from Philips Analytical, an earlier verson of X'Pert, has been used.

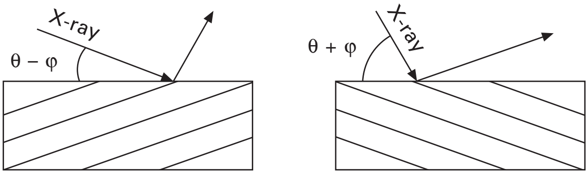
Unlike a fully strained layer, more effort must be spent on the structural characterization of a layer in which the lattice mismatch is partly relieved by misfit dislocations. Here, both the in-plane and the perpendicular layer parameters have to be determined, from which the residual strain and the Ge fraction can be obtained. It is required that at least two RCs have to be recorded and that at least one of them must stem from an asymmetric reflection, i.e. a from non-surface parallel (hkl) net plane. With θ being the Bragg angle and φ being the angle between the surface and the chosen net planes, two RCs can be obtained depending on whether the X-ray beam impinges on the (001) surface under the angle $\vartheta - \varphi$ (grazing incidence geometry) or under the angle $\vartheta + \varphi$ (grazing exit geometry) as illustrated in Fig. 1.11.

In the following these two schemes are denoted as (hkl)A or (hkl)B reflection, respectively. Of course, for a completely strain-free structure these two geometries reveal identical RCs. If, however, the layer is under strain, there are not only differences in the net plane distance Δd but also an angle difference $\Delta\varphi = \varphi_L - \varphi_S$ between substrate and layer as sketched in Fig. 1.12.

The consequence of this angle difference $\Delta\varphi$ for a RC is that the line spacing $\Delta\omega_A$ recorded in the (hkl)A reflection differs from $\Delta\omega_B$ recorded in the (hkl)B reflection. The angle differences $\Delta\omega_A$ and $\Delta\omega_B$ between the (hkl) Bragg peaks can be expressed as a function of the layer parameters f and $\varepsilon^{\|}$:

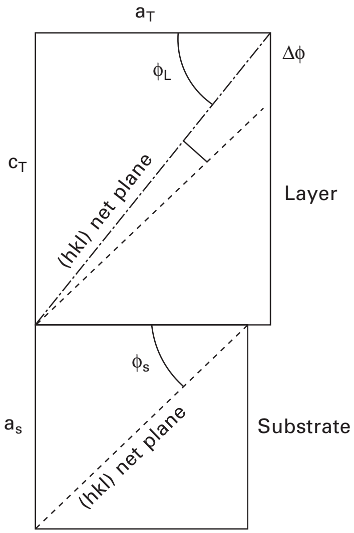
$$\Delta\omega_{A,B} = -\tan\vartheta\left[f + \varepsilon^{\|}\left(\sin^2\varphi - \frac{2\nu}{1-\nu}\cos^2\varphi\right)\right] \pm \frac{1}{2}\cdot\frac{1+\nu}{1-\nu}\varepsilon^{\|}\sin 2\varphi \qquad 1.20$$

where ϑ is the Bragg angle of the substrate and φ is the inclination angle between the diffracting (hkl) net planes and the (001) substrate surface



Grazing incidence Grazing exit

1.11 Two possible sample positions in the X-ray beam to record RCs on (hkl) net planes inclined by the angle φ, termed grazing incidence and grazing exit.



1.12 Schematic drawing of the same inclined (hkl) net planes of a strained layer and the (001) substrate showing the angle $\Delta\varphi$ caused by the tetragonal distortion of the layer lattice. L and S are the indices for layer and substrate, respectively.

according to the relation $\cos\phi = 1/(h^2 + k^2 + l^2)^{1/2}$. For an asymmetric reflection ($\phi = 0°$ and $\phi \neq 90°$) equation 1.20 has two solutions with $\Delta\omega_A$ for the + sign (grazing incidence) and $\Delta\omega_B$ for the – sign (grazing exit), respectively. By forming the sum and the difference of $\Delta\omega_A$ and $\Delta\omega_B$ gained from a pair of (hkl)A and (hkl)B RCs, the equations for $\varepsilon^{\|}$ and f read:

$$\varepsilon^{\|} = \frac{\Delta\omega_A - \Delta\omega_B}{[(1+\nu)/(1-\nu)]\sin 2\varphi} \tag{1.21}$$

$$f = -\frac{1}{2}(\Delta\omega_A + \Delta\omega_B)\cot\vartheta + \frac{1}{2}\left(\frac{1-\nu}{1+\nu}\right)\left(\frac{2\nu}{1-\nu}\cot\varphi - \tan\phi\right)(\Delta\omega_A - \Delta\omega_B) \tag{1.22}$$

These two relations enable a fast determination of layer strain and lattice mismatch, but they are restricted to an A/B pair of RCs taken from the same inclined (hkl) net planes.

Equation 1.20 can, however, also be used to derive more general equations for $\varepsilon^{\|}$ and f as functions of any ($h_ik_il_i$) and ($h_jk_jl_j$) RCs, where i and j are indices for different (hkl) indices. Not only A–B but also A–A and B–B combinations are possible. The equations read:

$$\varepsilon^{\|} = \Delta\omega_i \cot\vartheta_i - \Delta\omega_j \cot\vartheta_j$$

$$\times \left\{ \frac{1+\nu}{1-\nu} \left[(\cos^2\varphi_i - \cos^2\varphi_j) + \frac{1}{2}(u \cot\vartheta_i \sin 2\varphi_i - w \cot\vartheta_j \sin 2\varphi_j) \right] \Big] \right\}^{-1} \tag{1.23}$$

$$f = -\Delta\omega_i \cot\vartheta_i + \varepsilon^{\|} \left(\frac{1+\nu}{1-\nu} \right) \left[\cos\phi_i (u \cot\vartheta_i \sin\phi_i + \cos\phi_i) - \left(\frac{1-\nu}{1+\nu} \right) \right] \tag{1.24}$$

where

$\Delta\omega_i$ and $\Delta\omega_j$ = diffraction peak distances of the $(h_i k_i l_i)$A or B and $(h_j k_j l_j)$ A or B RCs

ϑ_{ij} = substrate Bragg angles for the $(h_i k_i l_i)$ and $(h_j k_j l_j)$ reflections

φ_{ij} = angle of the $(h_i k_i l_i)$ and $(h_j k_j l_j)$ net planes to the substrate surface

$u = 1$ and $u = -1$ = for $(h_i k_i l_i)$A and B reflection, respectively

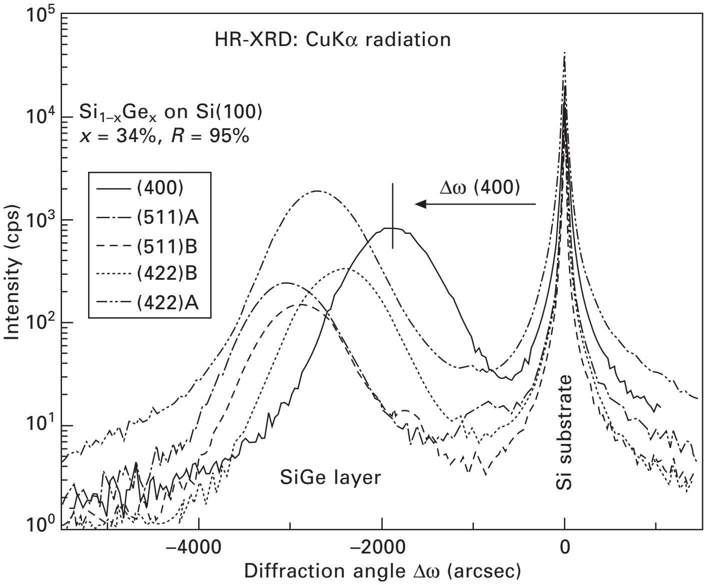
$w = 1$ and $w = -1$ = for $(h_j k_j l_j)$A and B reflection, respectively.

The derivation of relations 1.20–1.24, and also more details and some accuracy and difficulty aspects concerning the presented evaluation procedure, are given in [33].

As an example of a RC analysis from a strain-relieved SiGe layer on Si(100), five different (hkl) reflections are shown in Fig. 1.13. After graphical measurement of the line spacings, relations 1.23 and 1.24 are taken to determine the layer parameters $\varepsilon^{\|}$ and f. These values are then transferred in the Ge fraction of 34% and the degree of relaxation by using the relation $R = 1 - |\varepsilon^{\|}/f|$. For the Poisson ratio ν a value of 0.275 is assumed. The incoherent scattering contributions from the strain-relieving dislocations lead to a significant line broadening and the loss of thickness fringes. From the width β of the layer lines (corrected by the intrinsic line width) an estimation for the dislocation density ρ can be obtained by the simple relation $\rho = \beta^2/9b^2$ where b is the Burgers vector.

Finally, for example by a substantial deviation from the equipartition of the misfit dislocations, a layer tilt emerges, i.e. a small angle between the layer and substrate lattice normal. Such a tilt can also be measured quantitatively by recording a set of RCs under different sample azimuth angles ranging from 0° to 360°. From the sinus-like variation of the $\Delta\omega$ values the tilt angle can be simply determined.

Compared with the method of combining only pairs of (hkl)A and B measurements, the advantage of the free choice of reflection combinations is obvious. A measurement of n RCs enables $n(n-1)/2$ sets of values of in-plane strain $\varepsilon^{\|}$ and lattice mismatch f to be determined. Just as the exact determination of line spacing is often difficult because of the broad diffraction lines, the redundancy of this procedure gives a more accurate result.



1.13 Plot of five different (hkl) reflections recorded from a strain-relieved SiGe layer on Si(100) substrate. The evaluation reveals a $Si_{0.66}Ge_{0.34}$ layer on Si(100) with a degree of relaxation of 95%.

1.7 Future trends

The heterostructure SiGe/Si has contributed to a large extent to an understanding of lattice mismatched heterostructures and this understanding has led to rapidly increasing exploitation of SiGe in modern microelectronics. The SiGe heterobipolartransistor (HBT) has completely replaced earlier purely Si versions in high-speed bipolar and BiCMOS circuits. But the most widespread usage of SiGe, completely underestimated by the public, is its use in CMOS computer chips as a p-channel stressor via source/drain epitaxy. Applications in production now concentrate on low Ge content alloys because with higher mismatch surface corrugations increase. In future [34] structural investigations will be focused on:

- High Ge content alloys and possible ordering therein
- Templated self-organization of quantum dots [35]
- Non-equilibrium alloys (SiGe: C, Sn)
- Strain adjustment platforms such as virtual substrates or Ge On Insulator (GeOI) on large-diameter Si substrates.

1.8 Acknowledgement

The technical assistance of Mr. H. Xu in preparing this chapter and figure artwork is kindly achnowledged.

1.9 References

[1] Landolt-Börnstein, Vol. III/41 – *Condensed Matter, Semiconductors*, Subvolume A1 (Group IV Elements, IV–IV and III–V Compounds), Springer, Berlin (2001)

[2] Landolt-Börnstein, Vol. III/34 – *Condensed Matter, Semiconductors Quantum Structures*, Springer, Berlin (2001)

[3] *Properties of Crystalline Silicon*, EMIS Datareview Series, vol. 20, IEE INSPEC, London (1999)

[4] *Properties of Silicon Germanium and SiGe: Carbon*, EMIS Datareview Series, vol. 24, IEE INSPEC, London (2000)

[5] J.P. Dismukes, L. Ekstrom, R.J. Paff, *J. Phys. Chem.* 68, 3021–3027 (1964)

[6] E. Kasper, A. Schuh, G. Bauer, B. Hollander, H. Kibbel, *J. Cryst. Growth* 157, 68–72 (1995)

[7] R. Fabbri, F. Cembali, M. Servidori, A. Zani, *J. Appl. Phys.* 74, 2359–2369 (1993)

[8] S. de Gironcoli, P. Giannozzi, *Phys. Rev. Lett.* 66, 2116–2119 (1991)

[9] The relation is implemented in the simulation software X'Pert Epitaxy from PANalytical, http://www.panalytical.com/index.cfm?pid=354

[10] D. Windisch, P. Becker, *Phys. Stat. Sol.A* 118, 379–388 (1990)

[11] O. Brümmer, V. Alex, G. Schulze, *Ann. Phys.* 28, 118–134 (1972)

[12] J.F.C. Baker, M. Hart, *Acta Crystallogr. A* 31, 364–367, (1975)

[13] W.L. Bragg, *Phil. Mag.* 40, 169 (1920)

[14] L. Pauling, M.L. Huggins, *Z. Kristallogr. Kristallgeom. Kristallphys. Kristallchem.* (Germany) 87, 205 (1934)

[15] L. Vegard, *Z. Physik* (Germany) 5, 17–26 (1920)

[16] J.C. Aubry, T. Tyliszczak, A.P. Hitchcock, J.-M. Baribeau, T.E. Jackmann, *Phys. Rev. B* 59, 12872–12883 (1999)

[17] M. Mendik, M. Ospelt, H. von Känel, P. Wachter, *Appl. Surf. Sci.* 50, 303 (1991)

[18] P.X. Zhang, D.J. Lockwood, H.J. Labbe, J.-M. Baribeau, *Phys. Rev. B* 46, 9881–9884 (1992)

[19] H. Stohr, W. Klemm, *Z. Anorg. AIlg. Chem.* (Germany) 241, 305–323 (1939)

[20] F.X. Hassion, A.J. Goss, F.A. Trumbore, *J. Phys. Chem.* 59, 1118-1119 (1955)

[21] C.D. Thurmond, *J. Phys. Chem.* 57, 827–830 (1953)

[22] R.W. Olesinski, G.J. Abbaschian, *Bull. Alloy Phase Diagrams (USA)* 5, 180–183 (1984)

[23] J. Schilz, V.N. Romanenko, *J. Mater. Sci., Mater. Electron.* 6, 265–279 (1995)

[24] K. Alberi, J. Blacksberg, L.D. Bell, S. Nikzad, K.M. Yu, O.D. Dubon, W. Walukiewicz, *Phys. Rev. B* 77, 073202 (2008)

[25] S. Takeuchi, A. Sakai, K. Yamamoto, O. Nakatsuka, M. Ogawa, S. Zaima, *Semicond. Sci. Technol.* 22, S231–S235 (2007)

[26] J. Mathews, R. Roucka, J. Xie, S.-Q. Yu, J. Menéndez, J. Kouvetakis, *Appl. Phys. Lett.* 95, 133506 (2009)

[27] N. Thompson, *Proc. Phys. Soc. B* 66, 481 (1953)

[28] E. Kasper, H.J. Herzog, *Thin Solid Films* 44, 357 (1977)

[29] J.W. Matthews, A.E. Blakeslee, *J. Cryst. Growth* 27, 118 (1974)

[30] Q. People, J.C. Bean, *Appl. Phys. Lett.* 49, 229 (1986)

[31] Jordan Valley Semiconductor, *www.jvsemi.com*

[32] Bruker AXS, www.bruker-axs.com

[33] H.-J. Herzog, E. Kasper, *J. Cryst. Growth* 144, 177–186 (1994)
[34] E. Kasper, H.-J. Müssig, H.G. Grimmeiss (eds), *Advances in Electronic Materials* (Material Science Forum, vol. 608), TransTech Publications, Zürich (2009)
[35] http://www.nanosil-noe.eu/nanosil/

2

Electronic band structures of silicon–germanium (SiGe) alloys

N. MORI, Osaka University, Japan

Abstract: Electronic band structures of SiGe systems are described mainly from a theoretical point of view. The electronic band structures of bulk Si, bulk Ge, and SiGe alloys are briefly reviewed. Strain effects on the electronic band structures of pseudomorphic strained SiGe heterostructures are discussed in detail. Band offset and effective masses in those heterostructures are also presented.

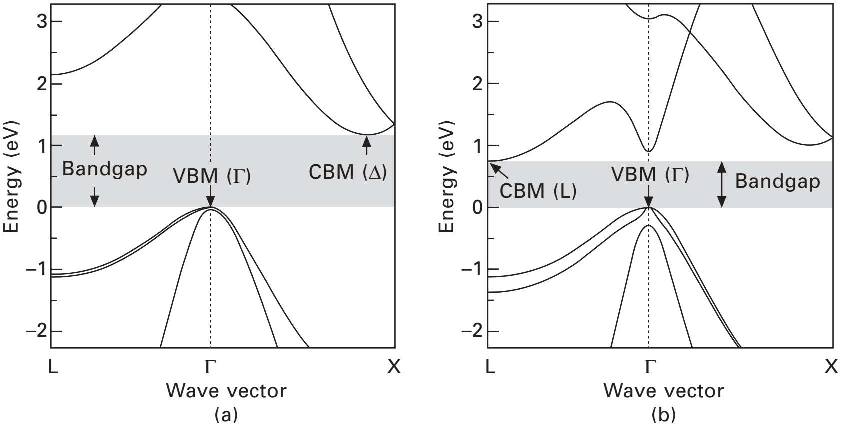
Key words: pseudomorphic SiGe, band structure, strain, band offset, effective mass.

2.1 Band structures

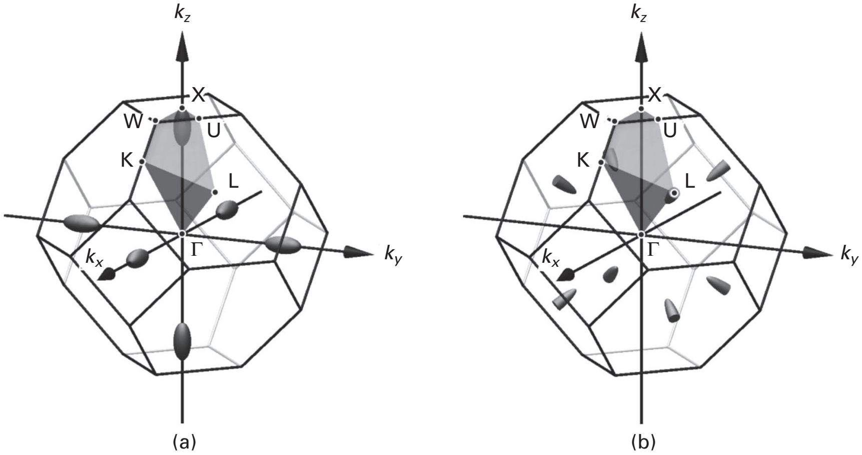
In this chapter, electronic band structures of SiGe systems are described mainly from a theoretical point of view. The electronic band structures of bulk Si, bulk Ge, and SiGe alloys are first reviewed (Yu and Cardona, 1999; Hamaguchi, 2010). Then, strain effects on the electronic band structures of pseudomorphic strained SiGe heterostructures are discussed. Finally, effective masses in those heterostructures are presented. The discussions in this chapter focus mainly on the most commonly studied SiGe heterostructures of (100)-strained $Si_{1-x}Ge_x$ on a Si substrate and (100)-strained Si on a relaxed virtual $Si_{1-y}Ge_y$ substrate. In the following, we include some numerical values of material parameters, which are cited from Madelung (2004) unless otherwise stated. Note, however, that widely scattered values of some material parameters have been reported and our choice of numerical values here is not comprehensive. For more comprehensive reviews, see, e.g., Fischetti and Laux (1996), Schaffler (1997), Brunner (2002), Cressler and Niu (2003), Yang *et al*. (2004), Paul (2004) and Lee *et al*. (2005).

2.1.1 Bulk silicon (Si) and germanium (Ge)

Si and Ge are both group IV elemental semiconductors and form the diamond crystal structure with a face-centered cubic arrangement of the atoms (see Fig. 1.1 of Chapter 1). The lattice constant of Ge is $a_{\mathrm{Ge}} = 0.5658$ nm, and is larger by 4.2% than the Si lattice constant, $a_{\mathrm{Si}} = 0.5431$ nm. There is a trend that semiconductors with a larger lattice constant are more likely to



2.1 Electronic band structure of (a) Si and (b) Ge calculated by the empirical tight-binding method (Jancu *et al.*, 1998).



2.2 Many-valley structure of (a) Si and (b) Ge.

have a smaller bandgap. This is the case for Si and Ge as shown in Fig. 2.1, in which the electronic band structures of Si and Ge are plotted. Both Si and Ge are indirect gap semiconductors. Si has the fundamental bandgap of E_g = 1.12 eV at T = 300 K. The conduction-band minima (CBM) of Si are located at the Δ point, 15% from the X point, along the ⟨100⟩ axes, and thus there are six equivalent conduction valleys in Si (see Fig. 2.2(a)). When a ⟨100⟩ axis is chosen as the principal axis of a newly defined ***k***-vector and the origins of ***k*** and E are set at the bottom of the conduction band, the dispersion relation is written as

$$E(\boldsymbol{k}) = \frac{\hbar^2}{2m_t}(k_x^2 + k_y^2) + \frac{\hbar^2}{2m_\ell}k_z^2 \qquad 2.1$$

For Si, the transverse effective mass is $m_t = 0.19m_0$ and the longitudinal effective mass is $m_\ell = 0.92m_0$, and Eq. 2.1 describes ellipsoidal constant energy surfaces as schematically shown in Fig. 2.2(a). Compared to Si, Ge has a significantly smaller bandgap of $E_g = 0.66$ eV at $T = 300$ K. The conduction-band minima of Ge are located at the L point along the $\langle 111 \rangle$ axes and there are four equivalent conduction valleys (or eight half-valleys) in Ge (see Fig. 2.2(b)). The E–$\boldsymbol{k}$ relation of the Ge conduction band is described also by Eq. 2.1 with $m_t = 0.82m_0$ and $m_\ell = 1.6m_0$. Unlike Si, the second conduction-band minimum, which is located at the Γ point, has relatively lower energy in Ge (see Fig. 2.1(b)) and can play a role in electron transport. The energy gap of this direct band-edge is $E_g = 0.81$ eV.

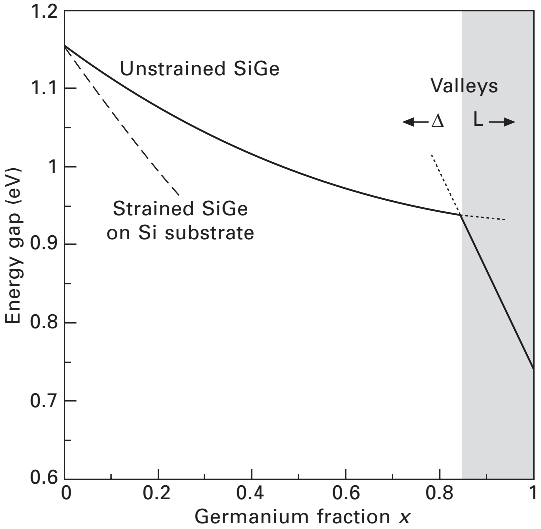
For the valence band, the valence-band maxima (VBM) is located at the Γ point in both Si and Ge (see Fig. 2.1). There are three principal valence bands of heavy-hole (HH), light-hole (LH), and split-off (SO) bands in Si and Ge. Ge has a significantly larger spin-orbit splitting ($\Delta_{Ge} = 0.30$ eV) compared to Si ($\Delta_{Si} = 0.044$ eV), which is primarily due to the heavier atom of Ge. The HH and LH bands are degenerated at the Γ point, so the interactions are strong for $\boldsymbol{k} \neq 0$ and the band shapes become non-parabolic and anisotropic, usually referred to as warped shapes. However, for simplicity, it is often expedient to assume that the valence-band effective masses are isotropic. In such a case, average hole masses are defined by averaging an anisotropic mass over all possible directions of $\boldsymbol{k}$. The averaged experimental hole masses are $m_{HH} = 0.54\mathrm{m}_0$, $m_{LH} = 0.15\mathrm{m}_0$, and $m_{SO} = 0.23\mathrm{m}_0$ for Si, and $m_{HH} = 0.34\mathrm{m}_0$, $m_{LH} = 0.043\mathrm{m}_0$, and $m_{SO} = 0.095\mathrm{m}_0$ for Ge (Cardona and Pollak, 1966; Yu and Cardona, 1999). We notice that Ge has smaller effective masses, especially for the LH band.

2.1.2 SiGe alloys

Si and Ge are miscible in all proportions, forming a continuous series of solid substitutional solutions of a fixed crystal structure over the entire composition range. The lattice constant $a_{SiGe}(x)$ of $Si_{1-x}Ge_x$ alloys shows minor departures from the linear dependence of Vegard's rule as discussed in Chapter 1. The fundamental bandgap of $Si_{1-x}Ge_x$ alloys is Si-like between the Γ-valence and Δ-conduction valleys up to a Ge content of about $x = 0.85$. The material has a Ge-like character with a conduction-band minimum at the L point for $x > 0.85$ (Braunstein *et al.*, 1958). Figure 2.3 shows the experimental low-temperature excitonic bandgap fitted by analytical expressions (Weber and Alonso, 1989):

$$E_{gx}^{\Delta}(x) = 1.155 - 0.43x + 0.206x^2 \text{ eV} \qquad \text{(Si-like)} \qquad 2.2$$

$$E_{gx}^{L}(x) = 2.010 - 1.270x \text{ eV} \qquad \text{(Ge-like)} \qquad 2.3$$



2.3 Low-temperature bandgap vs composition x for $Si_{1-x}Ge_x$ alloys (solid line), compared with the bandgap of a strained $Si_{1-x}Ge_x$ layer on a Si substrate (dashed line).

2.2 Strain effects

High-quality epitaxial interfaces can be produced between materials with different lattice constants. When the difference in lattice constants is small enough, the lattice mismatch is accommodated only by the tetragonal deformation of the epilayer. This so-called pseudomorphic layer is characterized by an in-plane lattice constant which remains the same throughout the structure. The strain present in such a structure dramatically alters its optical and electronic properties, in comparison to the bulk constituents. For example, the energy gap of strained $Si_{1-x}Ge_x$ on a Si substrate is substantially smaller than that of unstrained $Si_{1-x}Ge_x$ alloys as shown in Fig. 2.3, where the experimental low-temperature excitonic bandgap of strained $Si_{1-x}Ge_x$ fitted by an analytical expression (Robbins *et al.*, 1992)

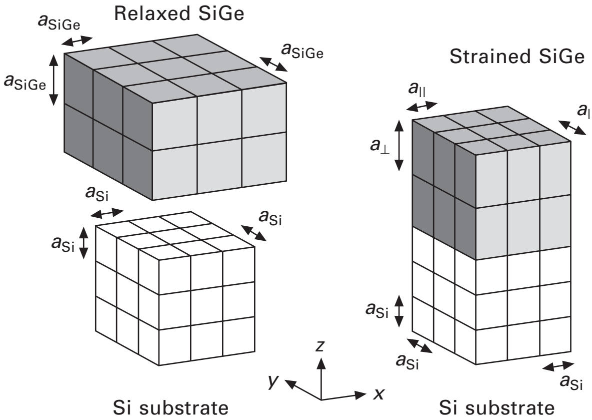
$$E_{gx}(x) = 1.155 - 0.874x + 0.376x^2 \text{ eV} \qquad (x < 0.24) \qquad 2.4$$

is plotted by the dashed line. In this section, strain effects are first described in detail for (100)-strained $Si_{1-x}Ge_x$ on a Si substrate (People, 1985; Van de Walle, 1989). The energy levels of (100)-strained $Si_{1-x}Ge_x$ layers on a $Si_{1-y}Ge_y$ substrate are then briefly discussed (Rieger and Vogl, 1993) followed by a discussion on the conduction and valence band offsets in strained-$Si_{1-x}Ge_x$/Si and strained-Si/$Si_{1-y}Ge_y$ heterostructures.

2.2.1 Strained $Si_{1-x}Ge_x$ on Si substrate

To understand strain effects on electronic band structures in strained $Si_{1-x}Ge_x$ on a Si substrate (see Fig. 2.4), we first describe the band alignment between unstrained $Si_{1-x}Ge_x$ and bulk Si by introducing an 'absolute' energy scale (Van de Walle, 1989). Then strain in the pseudomorphic $Si_{1-x}Ge_x$ layer is calculated within an elastic theory. After that, we discuss strain effects on the electronic band structure by calculating the energy shift and splitting of the band-edges of pseudomorphic $Si_{1-x}Ge_x$ layers on a Si substrate.

For heterostructures consisting of different semiconductors, the band lineups of the host semiconductors play an essential role in determining the electronic band structure. In principle it is possible to perform first-principles calculations of the band offsets at a semiconductor interface. However, a simple and reliable model to predict the band lineups is very useful for practical purposes. A theory based on the model-solid theory of Van de Walle and Martin gives us a simple and reliable means to examine not only lattice-matched but also strained layers (Van de Walle, 1989). The theory provides an absolute energy level for each semiconductor, which can be used to predict the band lineups of semiconductors. The absolute energy scale $E_{v,\mathrm{av}}$ is defined as an average of three uppermost valence bands at the Γ point ($E_{v,\mathrm{av}} = \frac{1}{3}(E_{\mathrm{HH}} + E_{\mathrm{LH}} + E_{\mathrm{SO}})$) and is parametrized in the literature (Van de Walle, 1989, Table II) for common IV and III–V semiconductors.



(a) (b)

2.4 (a) The lattice constant of a $Si_{1-x}Ge_x$ film to be grown on top of a Si substrate is larger than the Si lattice constant ($a_{\mathrm{SiGe}} > a_{Si}$). (b) The in-plane lattice constant remains the same throughout the structure for the pseudomorphic boundary condition ($a_{||} = a_{\mathrm{Si}} < a_{\mathrm{SiGe}}$). The biaxial compressive strain in the x–y plane leads to an extension along the z-direction ($a_{\perp} > a_{\mathrm{SiGe}}$).

The absolute energy scale $E_{v,\mathrm{av}}$ of Si is $E_{v,\mathrm{av}} = -7.03$ eV and that of Ge is $E_{v,\mathrm{av}} = -6.35$ eV. For alloys, linear interpolation between the pure materials is appropriate when considering a strain contribution (Cardona and Christensen, 1988; Van de Walle, 1989), since in a lattice mismatched alloy one material is effectively expanded, whereas the other is compressed. Rieger and Vogl (1993) have proposed a formula for $\mathrm{E}_{v,\mathrm{av}}$ differences in a general system of a strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ layer on a $\mathrm{Si}_{1-y}\mathrm{Ge}_y$ substrate:

$$\Delta E_{v,\mathrm{av}} = (0.47 - 0.06y)(x - y)\ \mathrm{eV} \tag{2.5}$$

where positive values refer to higher energies in the strained layer.

In a pseudomorphic system, the in-plane lattice constant $a_{\|}$ is the same throughout the structure. For $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ on a Si substrate, the in-plane lattice constant is thus equal to the bulk Si lattice constant ($a_{\|} = a_{\mathrm{Si}}$). Since the unstrained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ lattice constant is larger than the Si lattice constant ($a_{\mathrm{SiGe}} > a_{\mathrm{Si}}$), the strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ layer is subjected to biaxial compressive strain in the x–y plane perpendicular to the growth direction (z-direction). This leads to an extension normal to the plane so as to minimize the elastic energy. The lattice constant $a_{\perp}$ of the strained layer in the direction perpendicular to the interface is then given by

$$a_{\perp} = a_{\mathrm{SiGe}}[1 + D^{001}(1 - a_{\|}/a_{\mathrm{SiGe}})] \tag{2.6}$$

where the coefficient D^{001} is related to the elastic constants C_{11} and C_{12} as $D^{001} = 2C_{12}/C_{11}$ (see Table 2 of Chapter 1 for the numbers of C_{11} and C_{12}). The resulting strain tensor, which has only diagonal components, is given by

$$\varepsilon_{xx} = \varepsilon_{yy} = \frac{a_{\|} - a_{\mathrm{SiGe}}}{a_{\mathrm{SiGe}}} = \frac{a_{\mathrm{Si}} - a_{\mathrm{SiGe}}}{a_{\mathrm{SiGe}}} \tag{2.7}$$

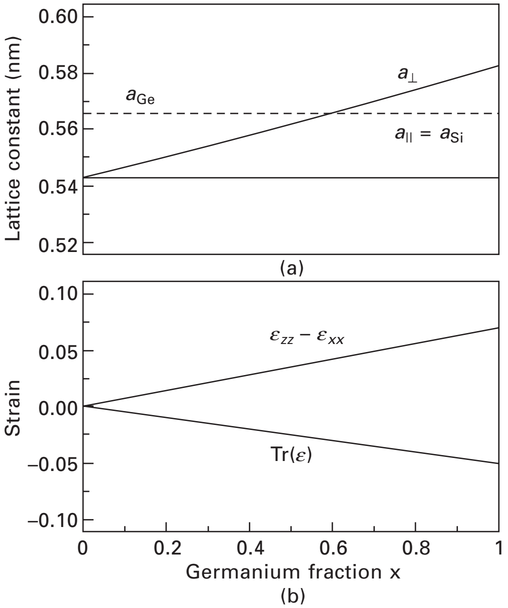
$$\varepsilon_{zz} = \frac{a_{\perp} - a_{\mathrm{SiGe}}}{a_{\mathrm{SiGe}}} = -D^{001}\varepsilon_{xx} \tag{2.8}$$

Figure 2.5 shows the lattice constants $a_{\|}$ and $a_{\perp}$ of a strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ layer, and the strain tensor components by the hydrostatic dilation $\Delta\Omega/\Omega = \mathrm{Tr}(\varepsilon) = \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz} = 2\varepsilon_{xx} + \varepsilon_{zz}$ and the deviatoric uniaxial component $\varepsilon_{zz} - \varepsilon_{xx} = (a_{\perp} - a_{\|})/a_{\mathrm{SiGe}}$.

The strain-induced energy change is originated in a hydrostatic component and a traceless deviatoric component. These components cause a shift and a splitting of degenerate band-edge levels, respectively (see Figs 2.6 and 2.7). The energy shift is proportional to the fractional volume change $\Delta\Omega/\Omega$ and is given by

$$\Delta E_{\mathrm{c}} = a_{\mathrm{c}}\mathrm{Tr}(\varepsilon) = a_{\mathrm{c}}(2\varepsilon_{xx} + \varepsilon_{zz}) \tag{2.9}$$

for the conduction band and



2.5 (a) Lattice constants $a_{||}$ and $a_\perp$ of a strained $\mathrm{Si}_{1-x}\,\mathrm{Ge}_x$ layer on a Si substrate. (b) Strain tensor components represented by the hydrostatic relative volume change Tr(ε) and the uniaxial component $\varepsilon_{zz} - \varepsilon_{xx}$.

$$\Delta E_{\mathrm{v}} = a_{\mathrm{v}}\mathrm{Tr}(\varepsilon) = a_{\mathrm{v}}(2\varepsilon_{xx} + \varepsilon_{zz}) \qquad 2.10$$

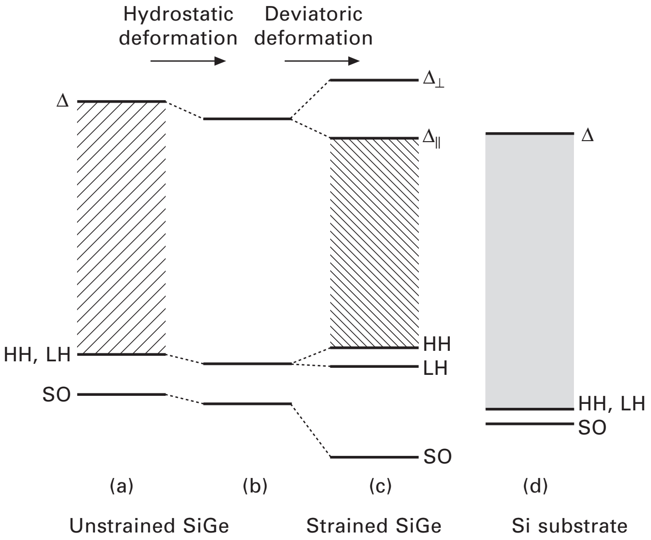
for the valence band, where a_{v} (a_{c}) is the dilation deformation potential for the valence (conduction) band. The bandgap deformation potential is equal to $a = a_{\mathrm{c}} - a_{\mathrm{v}}$. Their values are $a_{\mathrm{v}} = 2.46$ eV, $a_{\mathrm{c}}^{\Delta} = 4.18$ eV, and $a = 1.72$ eV, for Si, and $a_{\mathrm{v}} = 1.24$ eV, $a_{\mathrm{c}}^{\mathrm{L}} = -1.54$ eV, and $a = -2.78$ eV for Ge (Van de Walle, 1989).

The traceless part of the strain splits the six-fold degenerate Δ conduction valleys into four-fold $\Delta_{||}$ valleys within the layer plane and two-fold $\Delta_\perp$ valleys along the growth direction. The splitting of the bands is proportional to the uniaxial strain component, $\varepsilon_{zz} - \varepsilon_{xx}$, and is given by

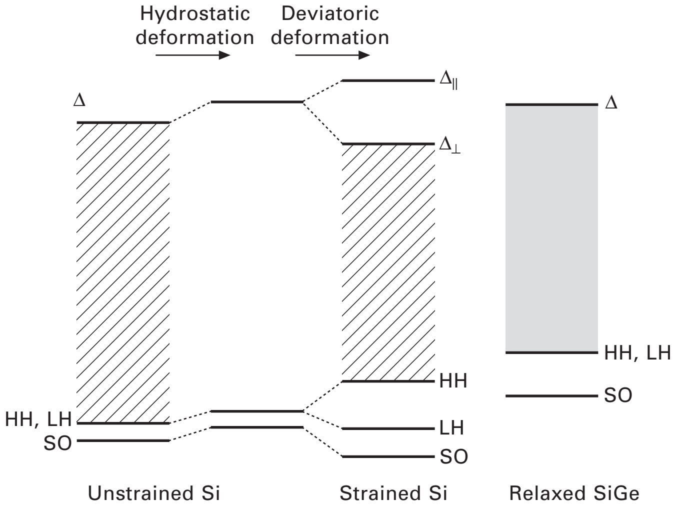
$$\Delta E_{\mathrm{c}}^{||} = -\frac{1}{3}\Xi_{\mathrm{u}}^{\Delta}(\varepsilon_{zz} - \varepsilon_{xx}) \qquad 2.11$$

$$\Delta E_{\mathrm{c}}^{\perp} = -\frac{2}{3}\Xi_{\mathrm{u}}^{\Delta}(\varepsilon_{zz} - \varepsilon_{xx}) \qquad 2.12$$

Since the deformation potential of Si is $\Xi_{\mathrm{u}}^{\Delta} = 8.7$ eV > 0 (Balslev, 1966), the four-fold $\Delta_{||}$ valleys move downward in energy and become lower than the two-fold $\Delta_\perp$ valleys for the biaxial compressive strain in a strained


Hydrostatic deformation
Deviatoric deformation
Δ
Δ⊥
Δ||
Δ
HH, LH
SO
HH
LH
SO
HH, LH
SO
(a)
(b)
(c)
(d)
Unstrained SiGe
Strained SiGe
Si substrate

2.6 Schematic band lineups in a strained SiGe layer on a Si substrate. (a) The valence band maximum (VBM) in a SiGe alloy is located above the VBM of a Si substrate. (b) The hydrostatic compressive strain reduces the band-edges. (c) The deviatoric strain components cause a splitting of degenerate band-edges. The heavy-hole (HH) valence band moves upward in energy, resulting in a large valence-band offset with respect to the Si substrate. (d) The four-fold $\Delta_{||}$ conduction valleys move downward with a quite small conduction band offset.


Hydrostatic deformation
Deviatoric deformation
Δ
Δ||
Δ⊥
Δ
HH, LH
SO
HH
LH
SO
HH, LH
SO
Unstrained Si
Strained Si
Relaxed SiGe

2.7 Schematic band lineups in a strained Si layer on a SiGe substrate.

SiGe layer on a Si substrate (see Fig. 2.6). For the biaxial tensile strain in a strained Si layer on a SiGe substrate, the two-fold $\Delta_\perp$ valleys become lower in energy than the four-fold $\Delta_\parallel$ valleys (see Fig. 2.7). Under the tetragonal deformation, the four L points remain equivalent and are conventionally labeled by N in a tetragonal lattice. The traceless part has no influence on their energy levels.

For valence bands of a strained SiGe alloy, the band structures can be described by the following 6 × 6 $\boldsymbol{k}\cdot\boldsymbol{p}$ Hamiltonian (we follow the notation of Chao and Chuang (1992)):

$$H(\boldsymbol{k}) = -\begin{bmatrix} P+Q & -S & R & 0 & -\frac{1}{\sqrt{2}}S & \sqrt{2}R \\ -S^\dagger & P-Q & 0 & R & -\sqrt{2}R & \sqrt{\frac{3}{2}}S \\ R^\dagger & 0 & P-Q & S & \sqrt{\frac{3}{2}}S^\dagger & \sqrt{2}Q \\ 0 & R^\dagger & S^\dagger & P+Q & -\sqrt{2}R^\dagger & -\frac{1}{\sqrt{2}}S^\dagger \\ -\frac{1}{\sqrt{2}}S^\dagger & -\sqrt{2}Q & \sqrt{\frac{3}{2}}S & -\sqrt{2}R & P+\Delta & 0 \\ \sqrt{2}R^\dagger & \sqrt{\frac{3}{2}}S^\dagger & \sqrt{2}Q & -\frac{1}{\sqrt{2}}S & 0 & P+\Delta \end{bmatrix} \tag{2.13}$$

with $P = P_k + P_\varepsilon$, $Q = Q_k + Q_\varepsilon$, $R = R_k + R_\varepsilon$, $S = S_k + S_\varepsilon$, and

$$P_k = \frac{\hbar^2}{2m_0}\gamma_1(k_x^2 + k_y^2 + k_z^2) \tag{2.14}$$

$$Q_k = \frac{\hbar^2}{2m_0}\gamma_2(k_x^2 + k_y^2 + 2k_z^2) \tag{2.15}$$

$$R_k = \frac{\hbar^2}{2m_0}\sqrt{3}[-\gamma_2(k_x^2 - k_y^2) + 2\mathrm{i}\gamma_3 k_x k_y) \tag{2.16}$$

$$S_k = \frac{\hbar^2}{2m_0}2\sqrt{3}\gamma_3(k_x - ik_y)k_z \tag{2.17}$$

$$P_\varepsilon = -a_\mathrm{v}\mathrm{Tr}(\varepsilon) \tag{2.18}$$

$$Q_\varepsilon = -\frac{b}{2}(\varepsilon_{xx} + \varepsilon_{yy} - 2\varepsilon_{zz}) \tag{2.19}$$

$$R_\varepsilon = \frac{\sqrt{3}}{2}b(\varepsilon_{xx} - \varepsilon_{yy}) - \mathrm{i}d\varepsilon_{xy} \tag{2.20}$$

$$S_\varepsilon = -d(\varepsilon_{xx} - \varepsilon_{yz}) \tag{2.21}$$

where Δ is the spin-orbit splitting, γ_1, γ_2, and γ_3 are the Luttinger parameters, and b and d are the Bir–Pikus shear deformation potentials. The band-edge levels can be obtained by setting $\boldsymbol{k} = 0$ in $\det[H(\boldsymbol{k}) - E] = 0$ and we have

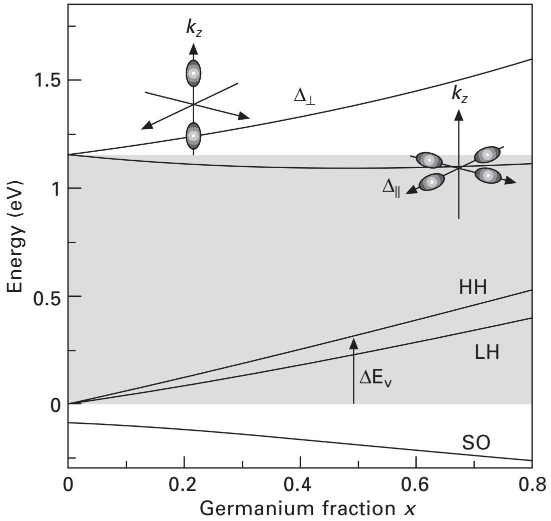
$$E_{\mathrm{HH}} = -P_\varepsilon - Q_\varepsilon \tag{2.22}$$

$$E_{\mathrm{LH}} = -P_\varepsilon - \frac{1}{2}[Q_\varepsilon - \Delta + \sqrt{\Delta^2 + 2\Delta Q_\varepsilon + 9Q_\varepsilon^2}\,] \tag{2.23}$$

$$E_{\mathrm{SO}} = -P_\varepsilon - \frac{1}{2}[Q_\varepsilon - \Delta - \sqrt{\Delta^2 + 2\Delta Q_\varepsilon + 9Q_\varepsilon^2}\,] \tag{2.24}$$

where use has been made of $\varepsilon_{xx} = \varepsilon_{yy}$ and $\varepsilon_{ij} = 0$ for $i \neq j$. The first terms on the right-hand sides of these equations ($-P_\varepsilon = a_{\mathrm{v}}\mathrm{Tr}(\varepsilon)$) correspond to the energy shift due to the hydrostatic strain component as already given in Eq. 2.10. The second terms, which depend on $-Q_\varepsilon = b(\varepsilon_{xx} - \varepsilon_{zz})$, represent effects of the deviatoric strain component and lift the degeneracy between the HH and LH bands.

Figure 2.8 shows a Ge fraction dependence of the band-edge levels of strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ on a Si substrate. The shaded region corresponds to the Si bandgap. The conduction-band minima consist of the four-fold $\Delta_{\|}$ valleys. The bandgap of strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ is significantly smaller than that of the unstrained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ alloy. The valence band offset, ΔE_{v}, is the most substantial in this system, while the conduction-band offset, ΔE_{c}, is quite small. Note that ΔE_{c} depends on the parameter set used for the calculation and can have the opposite sign to that shown in Fig. 2.8 (Rieger and Vogl, 1993).



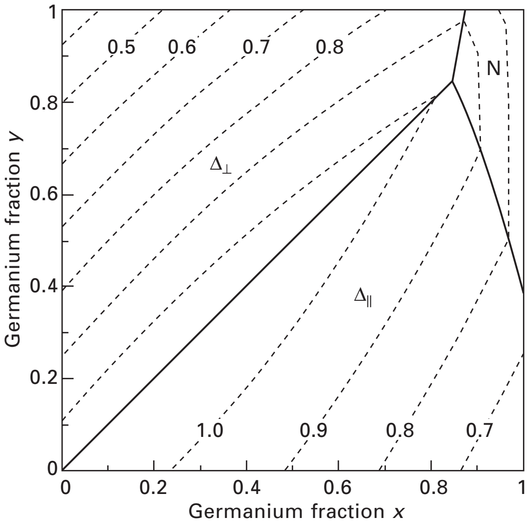
2.8 Band-edge levels of a strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ layer on a Si substrate as a function of Ge fraction x.

2.2.2 Strained $Si_{1-x}Ge_x$ on $Si_{1-y}Ge_y$ substrate

A complete set of band parameters for pseudomorphic (100)-strained $Si_{1-x}Ge_x$ on a $Si_{1-y}Ge_y$ substrate was presented by Rieger and Vogl (1993). The band parameters needed for device applications (bandgaps, band offsets, effective masses, and deformation potentials) are parametrized from non-local pseudopotential calculations with spin-orbit interaction. Figure 2.9 shows a contour plot of the fundamental bandgap of a strained $Si_{1-x}Ge_x$ layer on a $Si_{1-y}Ge_y$ substrate. The bandgap becomes largest in unstrained Si ($x = y = 0$) and smallest in a strained Si layer on a pure-Ge substrate ($x = 0$ and $y = 1$). When $x < y$, the strained $Si_{1-x}Ge_x$ layer is subjected to biaxial compressive strain because $a_{SiGe}(x) < a_{SiGe}(y)$. This results in the Si-like conduction-band minimum on the four-fold $\Delta_{\parallel}$ axis for most compositions. On the other hand, when $x > y$, the strained $Si_{1-x}Ge_x$ layer is subjected to biaxial tensile strain, resulting in the Si-like conduction-band minimum on the two-fold $\Delta_{\perp}$ axis. Only for weakly strained alloys with high Ge content does the conduction-band minimum lie at the N point, which corresponds to the L point for pure Ge.

2.2.3 Band offsets

The band alignments between heterointerfaces are one of the most important parameters for both transport and optical device applications. The precise



2.9 Fundamental bandgap in eV of a (100)-strained $Si_{1-x}Ge_x$ layer on a $Si_{1-y}Ge_y$ substrate as a function of x and y (Rieger and Vogl, 1993). The thick solid line represents the boundary between different conduction-band minima.

determination of the band offsets is, however, generally difficult in both experiments and theory.

A strained $Si_{1-x}Ge_x$ layer on a Si substrate has a large valence band offset ΔE_v for holes being confined in the strained $Si_{1-x}Ge_x$ layer (see Fig. 2.8). On the other hand, the conduction-band offset is very small, suggesting marginal type-I ($\Delta E_c = E_c(\mathrm{Si}) - E_c(Si_{1-x}Ge_x) > 0$) or type-II ($\Delta E_c < 0$) heterostructures. The experimental values for the band offsets are somewhat uncertain, but show a clear trend of a near-linear dependence of ΔE_v on a Ge fraction, x, for low x. The valence-band offset can be approximated by (Yang *et al.*, 2004)

$$\Delta E_v = 0.71x \text{ eV} \tag{2.25}$$

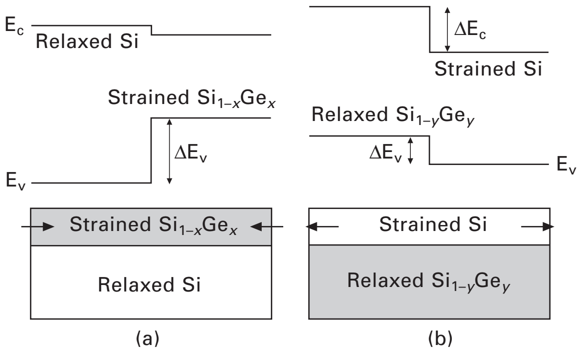
For a strained Si layer on a $Si_{1-y}Ge_y$ substrate, the system is a type-II heterostructure with larger band offsets for both the conduction and valence bands (see Fig. 2.10(b)). The valence-band and conduction-band offsets can be approximated by (Yang *et al.*, 2004)

$$\Delta E_v = 0.238y - 0.03y^2 \text{ eV} \tag{2.26}$$

and

$$\Delta E_c = 0.35y + 0.35y^2 - 0.12y^3 \text{ eV} \tag{2.27}$$

For a general system of a strained $Si_{1-x}Ge_x$ layer on a $Si_{1-y}Ge_y$ substrate, the valence band maximum occurs always in the layer with the higher Ge content (Schaffler, 1997). A type-I alignment is predicted for a Ge-rich strained layer on a Ge-rich substrate (Rieger and Vogl, 1993).



2.10 Schematic band alignments for (a) a biaxial compressively strained $Si_{1-x}Ge_x$ layer on a Si substrate and (b) a biaxial tensile strained Si layer on a $Si_{1-y}Ge_y$ substrate.

2.3 Effective mass

In this section, we first consider electron effective masses in strained SiGe layers. As discussed in the previous section, the six-fold degenerate Δ conduction valleys in unstrained Si split into the four-fold $\Delta_{\|}$ and the two-fold $\Delta_{\perp}$ conduction valleys in a strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ layer on a $\mathrm{Si}_{1-y}\mathrm{Ge}_y$ substrate. This results in a change in the electron effective mass. Let us consider a change in the conductivity mass, which is particularly important for transport applications, for example. For the six-fold Δ valleys, the conductivity mass can be written as

$$m_{\mathrm{c}} = \frac{3}{m_{\ell}^{-1} + 2m_{\mathrm{t}}^{-1}} \tag{2.28}$$

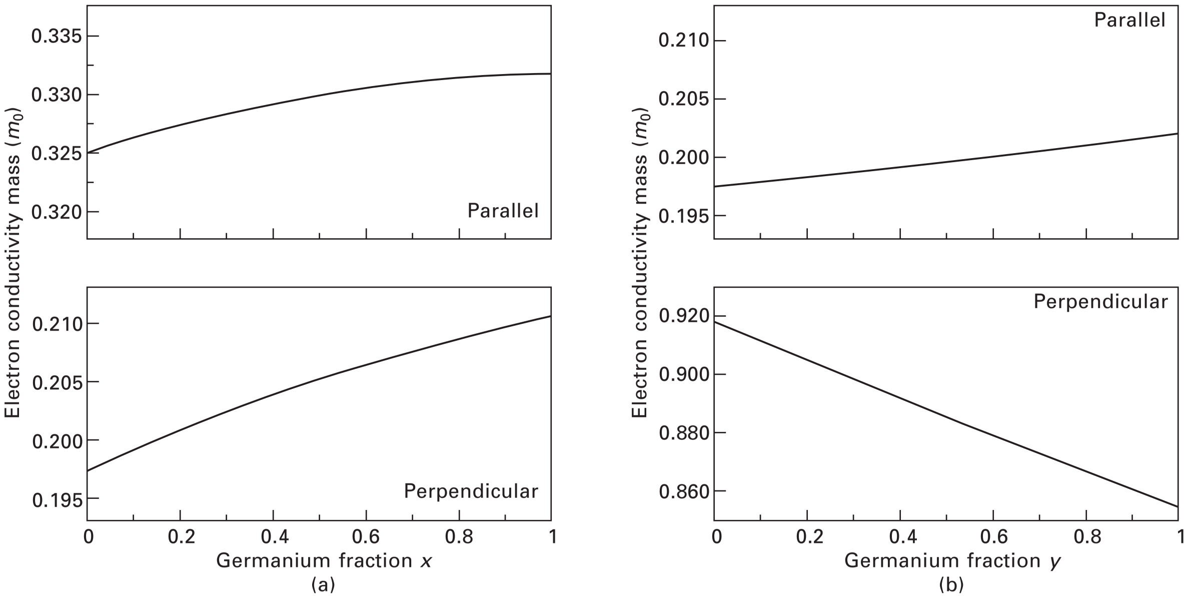
For the four-fold $\Delta_{\|}$ valleys, this will change into

$$m_{\mathrm{c}}^{\|} = \frac{2}{m_{\ell}^{-1} + m_{\mathrm{t}}^{-1}}, \quad m_{\mathrm{c}}^{\perp} = m_{\mathrm{t}} \tag{2.29}$$

in the in-plane (x–y) and perpendicular (z) directions, respectively. For the two-fold $\Delta_{\perp}$ valleys, $m_{\mathrm{c}}^{\|} = m_{\mathrm{t}}$ and $m_{\mathrm{c}}^{\perp} = m_{\ell}$. These types of changes in the effective mass, which originate in a change in the location of the conduction-band minimum, significantly affect the transport and optical properties of the SiGe heterostructures. A most significant change in the effective mass will occur when the L minimum sinks energetically below the Δ minimum for weakly strained alloys with high Ge content. Note, however, that we have to take into account all the energetically close lying conduction-band minima at finite temperatures, which will smooth out discontinuities in the effective masses associated with an abrupt change in the conduction-band minima.

In addition to the above-mentioned changes in the electron effective mass, the strain and Ge fraction cause a change in the effective mass value relative to the bulk Si value. Figure 2.11 shows a Ge fraction dependence of the conductivity mass in strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ on a Si substrate and in strained Si on a $\mathrm{Si}_{1-y}\mathrm{Ge}_y$ substrate (Rieger and Vogl, 1993). The effective masses weakly depend on the Ge fraction. For example, the parallel conductivity mass, $m_{\mathrm{c}}^{\|}(y)$, in strained Si on a $\mathrm{Si}_{1-y}\mathrm{Ge}_y$ substrate increases slightly with the Ge fraction y from the bulk Si value of m_{t} at $y = 0$.

We next consider hole effective masses in strained SiGe layers. A simple effective-mass description is not usually sufficient for describing the valence-band structure of strained SiGe heterostructures, since it is highly non-parabolic and anisotropic. Once an appropriate $\boldsymbol{k}\cdot\boldsymbol{p}$ band-parameter set is known, the six-band $\boldsymbol{k}\cdot\boldsymbol{p}$ Hamiltonian of Eq. 2.13 can be used for accurately describing the valence-band structure of strained SiGe alloys. Rieger and Vogl (1993) reported the $\boldsymbol{k}\cdot\boldsymbol{p}$ band parameters L, M, and N of the Dreselhaus notation,



2.11 Ge fraction dependence of the conductivity masses in (a) a strained $Si_{1-x}Ge_x$ layer on a Si substrate and (b) a strained Si layer on a $Si_{1-y}Ge_y$ substrate calculated from the parametrized effective masses of Rieger and Vogl (1993).

which fit remarkably well with the pseudopotential band structure, as (P standing equally for L, M, or N)

$$P(x) = P(0) + \alpha \ln(1 - Sx^{\beta}) \tag{2.30}$$

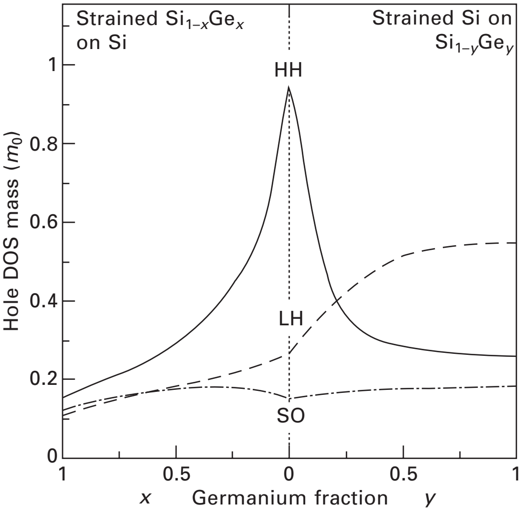
$$S = 1 - \exp\{[P(1) - P(0)]/\alpha\} \tag{2.31}$$

with $\alpha = 6.7064$, $\beta = 1.35$, $L(0) = -6.69$, $M(0) = -4.62$, $N(0) = -8.56$, $L(1) = -21.65$, $M(1) = -5.02$, and $N(1) = -23.48$ (units are $\hbar^2/2m_0$). The L, M, and N parameters can be converted to the Luttinger parameters through $\gamma_1 = -(L + 2M)/3 - 1$, $\gamma_2 = -(L - M)/6$, and $\gamma_3 = -N/6$, where L, M, and N are in units of $\hbar^2/2m_0$.

As an example of the valence-band structure calculation based on the six-band $\boldsymbol{k{\cdot}p}$ Hamiltonian, Fig. 2.12 shows a Ge fraction dependence of hole density-of-states (DOS) masses, m_d, at the room-temperature thermal energy $E_{th} = \frac{3}{2}kT$ in strained $Si_{1-x}Ge_x$ on a Si substrate and in strained Si on a $Si_{1-y}Ge_y$ substrate (Fischetti and Laux, 1996; Yang *et al.*, 2004). The DOS mass m_d is defined through

$$\frac{1}{2\pi^2}\left(\frac{2m_d}{\hbar^2}\right)^{3/2} E_{th}^{1/2} = 2\sum_{\boldsymbol{k}} \delta(E_{th} - E_{kin}(\boldsymbol{k})) \tag{2.32}$$

where $E_{kin}(\boldsymbol{k})$ $(= E(0) - E(\boldsymbol{k}))$ is the hole kinetic energy. The substantial reduction in m_d of the HH band (the band coming from the Si HH band) can be seen for a finite Ge fraction in both strained $Si_{1-x}Ge_x$ on a Si substrate



2.12 Ge fraction dependence of hole density-of-states masses at T = 300 K for a strained $Si_{1-x}Ge_x$ layer on a Si substrate (left half) and a strained Si layer on a $Si_{1-y}Ge_y$ substrate (right half).

and strained Si on a $Si_{1-y}Ge_y$ substrate. Note that the HH band is the topmost valence band in strained $Si_{1-x}Ge_x$ on a Si substrate, while the LH band is at the top of the valence band in strained Si on a $Si_{1-y}Ge_y$ substrate (see Figs 2.6 and 2.7).

2.4 Conclusion

In a pseudomorphic SiGe layer, the lattice mismatch is accommodated by the tetragonal deformation of the layer. The strain present in such a layer dramatically alters its optical and electronic properties, in comparison to the bulk constituents. The strain effects on the electronic band structures can be understood by considering (1) the band alignment between unstrained host semiconductors, (2) strain present in a pseudomorphic layer, which can be estimated from an elastic theory, and (3) the energy shift and splitting of the band-edges due to the strain. The fundamental bandgap of a strained $Si_{1-x}Ge_x$ layer on a $Si_{1-y}Ge_y$ substrate becomes largest in unstrained Si ($x = y = 0$) and smallest in a strained Si layer on a pure-Ge substrate ($x = 0$ and $y = 1$). Effective mass variations originating from a change in the location of the band extrema have a strong impact on the transport properties of the SiGe heterostructures.

2.5 References

Balslev I (1966), 'Influence of uniaxial stress on indirect absorption edge in silicon and germanium', *Physical Review*, 143, 636–647.

Braunstein R, Moore AR and Herman F (1958), 'Intrinsic optical absorption in germanium–silicon alloys', *Physical Review*, 109, 695–710.

Brunner K (2002) , 'Si/Ge nanostructures', *Reports on Progress in Physics*, 65, 27–72.

Cardona M and Christensen NE (1988), 'Comment on "Spectroscopy of excited states in $In_{0.53}Ga_{0.47}As$-InP single quantum wells grown by chemical-beam epitaxy" ', *Physical Review B*, 37, 1011–1012.

Cardona M and Pollak FH (1966), 'Energy-band structure of germanium and silicon: The k·p method', *Physical Review*, 142, 530–543.

Chao CYP and Chuang SL (1992), 'Spin-orbit-coupling effects on the valence-band structure of strained semiconductor quantum-wells', *Physical Review B*, 46, 4110–4122.

Cressler JD and Niu G (2003), *Silicon–Germanium Heterojunction Bipolar Transistors*, Boston, MA, Artech House.

Fischetti MV and Laux SE (1996), 'Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys', *Journal of Applied Physics*, 80, 2234–2252.

Hamaguchi C (2010), *Basic Semiconductor Physics*, Berlin, Splinger-Verlag.

Jancu JM, Scholz R, Beltram F and Bassani F (1998), 'Empirical spds* tight-binding calculation for cubic semiconductors: General method and material parameters', *Physical Review B*, 57, 6493–6507.

Lee ML, Fitzgerald EA, Bulsara MT, Currie MT and Lochtefeld A (2005), 'Strained

Si, SiGe, and Ge channels for high-mobility metal-oxide-semiconductor field-effect transistors', *Journal of Applied Physics*, 97, 011101 (1–27).

Madelung O (2004), *Semiconductors: Data Handbook*, Berlin, Springer-Verlag.

Paul DJ (2004), 'Si/SiGe heterostructures: from material and physics to devices and circuits', *Semiconductor Science and Technology*, 19, R75–R108.

People R (1985), 'Indirect band-gap of coherently strained Ge_xSi_{1-x} bulk alloys on (001) silicon substrates', *Physical Review B*, 32, 1405–1408; Erratum: *ibid.*, 33, 1451 (1986).

Rieger MM and Vogl P (1993), 'Electronic-band parameters in strained $Si_{1-x}Ge_x$ alloys on $Si_{1-y}Ge_y$ substrates', *Physical Review B*, 48, 14276–14287; Erratum: *ibid.*, 50, 8183 (1994).

Robbins DJ, Canham LT, Barnett SJ, Pitt AD and Calcott P (1992), 'Near-band-gap photoluminescence from pseudomorphic $Si_{1-x}Ge_x$ single layers on silicon', *Journal of Applied Physics*, 71, 1407–1414.

Schaffler F (1997), 'High-mobility Si and Ge structures', *Semiconductor Science and Technology*, 12, 1515–1549.

Van de Walle CG (1989), 'Band lineups and deformation potentials in the model-solid theory', *Physical Review B*, 39, 1871–1883.

Weber J and Alonso MI (1989), 'Near-band-gap photoluminescence of Si-Ge alloys', *Physical Review B*, 40, 5683–5693.

Yang LF, Watling JR, Wilkins RCW, Borici M, Barker JR, Asenov A and Roy S (2004), 'Si/SiGe heterostructure parameters for device simulations', *Semiconductor Science and Technology*, 19, 1174–1182.

Yu PY and Cardona M (1999), *Fundamentals of Semiconductors*, Berlin, Springer-Verlag.

3

Understanding crystal growth mechanisms in silicon–germanium (SiGe) nanostructures

M. SUEMITSU, Tohoku University, Japan and
S. N. FILIMONOV, Tomsk State University, Russia

Abstract: This chapter recalls some basic concepts required for understanding crystal growth. Thermodynamics and kinetics of the crystallization process are considered and an atomistic picture of the fundamental surface processes is given. Basic mechanisms of kinetic growth instabilities in epitaxial growth are described. Finally, an overview of the strain relaxation mechanism in heteroepitaxial growth is given using the Ge/Si(111) and Ge/Si(001) systems as an example.

Key words: supersaturation, surface diffusion, island nucleation, kinetic growth instabilities, strain relaxation in heteroepitaxial growth.

3.1 Introduction

Complementary metal oxide semiconductor (CMOS)-based silicon technology is facing its fundamental limitations. The scaling strategy led by Moore's law, with which both high performance and high integration have been simultaneously achieved, is now approaching a point at which increasing power consumption can hardly be paid back. New strategies other than scaling are strongly demanded, which include *More Moore*, in which new gate and channel materials or strain are introduced to the devices, *More than Moore*, in which new functionalities such as sensors and micro mechanical systems (MEMS) are attached to the conventional CMOS devices, and *Beyond CMOS*, in which novel operating principles such as quantum effects and spintronics are being considered.

Breakthroughs to the fundamental limitations of silicon are being realized by introducing new materials with nanostructures. Our ability to grow high-quality thin crystalline films of new materials and to control their shape and size on the nanometer scale are therefore key to the next generation of silicon-based technology.

In the semiconductor industry top-down methodologies based on optical lithography, ion and electron beams are common practice. However, limitations of top-down methodologies can be found in their spatial resolution (optical lithography) and in their throughputs (beam-based methods), which demand researchers to seek bottom-up methodologies, i.e. self-organized patterning, in the device fabrication processes. In this respect, crystal growth, again, can be very instructive.

The major objective of this chapter is to recall fundamentals of crystal growth, in order to give readers some key concepts which are required for understanding and constructing bottom-up nanofabrication processes. For a deeper acquaintance with crystal growth fundamentals the reader is referred to the classical monograph by A. A. Chernov (Chernov, 1984) and to some recent textbooks on crystal growth (Pimpinelli and Villain, 1998; Markov, 2003; Michely and Krug, 2003).

This chapter consists of five sections. In Section 3.2, we recall the basic thermodynamic concepts, relevant to crystal growth, such as surface and interfacial energies, chemical potential, and supersaturation. Using these concepts, a classification of the equilibrium growth modes (Volmer–Weber, Frank–van-der-Merwe, Stranski–Krastanov) is introduced. Section 3.3 deals with atomic-scale aspects of crystal growth. We start the discussion by introducing the very basic concepts of adsorption and desorption, together with their reaction kinetics. As an example of interplay between adsorption and desorption, the surface chemistry of Si gas-source MBE is briefly introduced. Then we proceed to discussing the atomistic aspects of crystal growth, by introducing the key structural elements such as surface steps, kinks and adatoms as the main players of the game. Based on them, surface processes like surface diffusion, adatom incorporation, step crossing, and island nucleation are discussed. Section 3.4 describes growth at far-from-equilibrium conditions. The section covers such topics as step-flow and layer-by-layer growth modes, nucleation kinetics and 2D island formation, and kinetic growth instabilities (step bunching, step meandering, mound formation). Finally, in Section 3.5 we discuss the strain-induced effects in epitaxial growth. The concept of misfit strain is introduced and basic mechanisms of the strain relaxation are discussed. An interplay between thermodynamics and kinetics in the nucleation and growth of strained 3D islands is discussed using the Ge/Si(111) and Ge/Si(001) systems as an example.

3.2 Thermodynamics of crystal growth

During crystal growth the growing species (atoms or molecules) flow from an ambient phase into the crystalline phase. The ambient phase may be a gas, a liquid or another solid, depending on the applied method of crystal growth. We will consider first infinite (bulk) phases. The thermodynamic quantity that determines the relative stability of the participating phases is the chemical potential. The chemical potential of a given phase is the work that has to be done to a single particle (atom or molecule) to be incorporated into this phase (to change the number of particles in this phase by unity). If two or more phases are in contact, the phase with the lowest chemical potential is the most stable one. This phase will grow at the expense of other phases having higher chemical potentials. If the chemical potentials of the ambient phase μ_{amb} and growing crystal μ_{C} are equal, the phases are in equilibrium and no phase transition occurs. The difference

$$\Delta\mu = \mu_{\mathrm{amb}}(P, T) - \mu_{\mathrm{C}}(P, T) \tag{3.1}$$

if any, is called supersaturation, and represents the measure of the driving force for the phase transition. Generally, larger absolute values of $\Delta\mu$ correspond to higher rates of the phase transition. If $\Delta\mu$ is positive, the crystal will grow. If $\Delta\mu$ is negative, the reverse process must take place.

For practical use it is convenient to express the supersaturation $\Delta\mu$ in terms of the parameters that can be directly handled in the experiment, e.g. pressure P and temperature T. For instance, if the ambient phase is an ideal gas, its chemical potential is defined as $\mu_{\mathrm{V}} = \mu_{\mathrm{V}}^{0} + k_{\mathrm{B}}T \ln P$, where the standard chemical potential μ_{V}^{0} is the Gibbs energy per atom of the pure ideal gas at the standard pressure of 1 bar, and k_{B} is the Boltzmann constant. If the system is kept at a constant temperature T, the supersaturation $\Delta\mu$ can be written as

$$\Delta\mu = k_{\mathrm{B}}T \ln \frac{P}{P_0} \tag{3.2}$$

where P_0 is the equilibrium vapor pressure at temperature T. If the deviation $\Delta P = P - P_0$ from the equilibrium pressure is small, then Eq. 3.2 is reduced to $\Delta\mu = k_{\mathrm{B}}T\Delta P/P_0$. The difference ΔP is called the absolute supersaturation or the absolute overpressure and the ratio $\sigma \equiv \Delta P/P_0$ defines the relative supersaturation.

A very similar expression for $\Delta\mu$ is obtained for crystallization from a binary solution. Let C and C_0 be the real and the equilibrium concentrations of the solute, respectively. Assuming that only solute atoms incorporate into the crystal, and treating the solution as an ideal solution ($\mu_{\mathrm{L}} = \mu_{\mathrm{L}}^{0} + k_{\mathrm{B}}T \ln C$), we get

$$\Delta\mu = k_{\mathrm{B}}T \ln \frac{C}{C_0} \tag{3.3}$$

which is reduced to $\Delta\mu = k_{\mathrm{B}}T\Delta C/C_0$ if the supersaturation $\Delta C = C - C_0$ is small. The ratio $\sigma = \Delta C/C_0$ determines the relative supersaturation.

To calculate the supersaturation $\Delta\mu$ in a crystal–melt system we note that for a single-component system the chemical potential of a given phase is equal to the Gibbs free energy per atom of that phase, therefore $\mathrm{d}\mu = -s\mathrm{d}T + v\mathrm{d}P$. If the melt is cooled down below the equilibrium temperature T_0 by ΔT under a constant pressure, the supersaturation $\Delta\mu$ will be given by

$$\Delta\mu = (s_L - s_C)\Delta T \equiv \Delta h \frac{\Delta T}{T_0} \qquad 3.4$$

Here s_L and s_C are the entropy per atom in the liquid phase and in the crystal, respectively, and $\Delta h = (s_L - s_C)\,T_0$ is the enthalpy (heat) of melting at the melting temperature T_0. The difference $\Delta T = T_0 - T$ is called undercooling, and the relative supersaturation is given by $\sigma \equiv \Delta T/T_0$.

So far, we have been speaking of chemical potentials of infinitely large (bulk) phases. It is clear, however, that a real growing crystal has a finite size, and this size can be very small, e.g. at the very beginning of the crystallization process. What makes a finite crystal so different from the infinite one is the increased role of the surface, a boundary that separates the crystal from the ambient. Since atomic bonds are partially broken at the surface, creation of a new surface costs energy. The minimum work α_S required to create a unit surface area under a constant volume and temperature is called the surface free energy. If we take a crystal of spherical shape with radius R and add one atom with volume Ω to that crystal, its surface area must increase by $2\Omega/R$. Therefore, an extra work $2\Omega\alpha_S/R$ has to be performed to increase the size of the spherical crystallite by one atom. Hence, the chemical potential of a spherical crystallite is larger by this amount than the chemical potential of the bulk:

$$\mu_C = \mu_C^{\infty} + \frac{2\Omega\alpha_S}{R} \qquad 3.5$$

Here we put the superscript ∞ to distinguish the chemical potential of the bulk crystal. This equation, known as the Gibbs–Thomson equation, tells us that smaller crystallites with a higher surface curvature are thermodynamically less stable than larger ones.

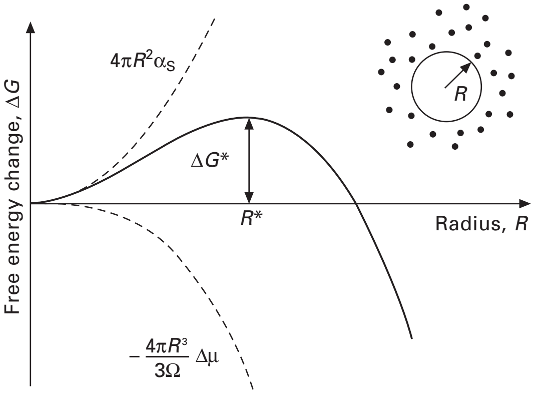
Due to the surface energy contribution, the nucleation of a crystal from a dilute phase is a thermally activated process. To see this, let us consider the change of the Gibbs free energy of the system due to formation of a spherical crystallite of radius R:

$$\Delta G = -\,i\Delta\mu + 4\pi R^2\alpha_S \qquad 3.6$$

The first term on the right-hand side is the energy gained when $i = 4\pi R^3/3\Omega$ atoms from the ambient phase join the crystal (here, as before, $\Delta\mu = \mu_{\text{amb}} - \mu_C^{\infty}$ is referred to the supersaturation between two bulk phases) and the second term is the surface free energy of the crystallite. The $\Delta G(R)$ dependence plotted in Fig. 3.1 shows a maximum at some critical radius R^* given by

$$R^* = \frac{2\alpha_S\Omega}{\Delta\mu} \qquad 3.7$$

The crystallites smaller than R^* are unstable and decay with a high probability,



3.1 Change of the Gibbs free energy due to the formation of a spherical crystallite.

while those larger than R^* may decrease their free energy by getting more atoms from the ambient phase and therefore should keep growing. The cluster of the critical radius R^* is called the critical nucleus.

The maximal value of ΔG determines the free energy barrier, ΔG^*, which has to be surmounted to form the critical nucleus. Substituting Eq. 3.7 into Eq. 3.6 one gets

$$\Delta G^* = \frac{4}{3}\pi R^{*2}\alpha_S = \frac{16\pi}{3}\frac{\alpha_S^2\Omega^2}{\Delta\mu^2} \qquad 3.8$$

It follows from Eq. 3.8 that the nucleation barrier decreases rapidly with increasing supersaturation $\Delta\mu$.

When a crystal is grown on a substrate, one more contribution to the system's free energy arises. It is the free energy of the film–substrate interface, or interfacial energy per unit area, α_{int}, which is defined as the work required to increase the interface by unit area. This quantity can be viewed as a measure of the bond strength between the film and the substrate atoms, with weaker bonds corresponding to a higher interfacial energy and vice versa.

When a film is growing under close to equilibrium conditions, mutual strength among the surface free energy of the substrate α_S, the surface free energy of the film α_F, and the interfacial free energy α_{int} determines the growth mode and the morphology of the growing film. Three major growth modes are generally distinguished. When the surface free energy of the substrate is lower than the sum of the free energies for the film and the interface:

$$\alpha_S < \alpha_F + \alpha_{\text{int}} \qquad 3.9$$

it is energetically favorable to keep as much uncovered substrate surface area as possible. This is achieved by the formation of three-dimensional (3D) islands directly on the substrate surface. This growth mode is known as the Volmer–Weber (VW) mode (Fig. 3.2(a)).

In the opposite case of a high surface energy of the substrate:

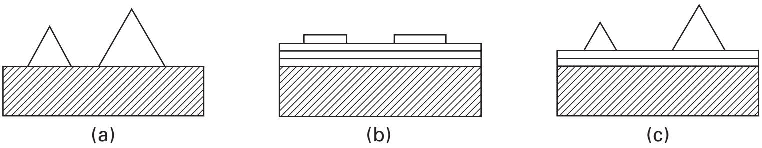
$$\alpha_S \geq \alpha_F + \alpha_{int} \qquad 3.10$$

the growing film tends to cover the substrate completely to eliminate the contribution of the high substrate surface energy α_S. This leads to the layer-by-layer growth mode, also known as Frank–van-der-Merwe (FM) growth (Fig. 3.2(b)). FM growth is predicted also for homoepitaxial growth, where the film and the substrate are of the same material. In this case $\alpha_S = \alpha_F$ and $\alpha_{int} = 0$ by definition, and the Gibbs–Thomson effect keeps the growing surface flat.

When a crystalline film is grown on a substrate of a different material, there might be a non-negligible lattice mismatch between the two crystalline structures. In this case the film surface might become three-dimensional after an initial layer-by-layer growth for a few layers (Fig. 3.2(c)). This scenario, known as the Stranski–Krastanov (SK) growth mode, is typical for growth of Ge films on Si surfaces. In Stranski–Krastanov growth the system starts with condition 3.10, which leads to layer-by-layer growth. As the film gets thicker, the chemical influence of the substrate diminishes, and the two surface energies α_S and α_F become essentially the same. At the same time, elastic strain accumulates in the growing film, leading to a weaker adhesion of the depositing material to the already completed layers. The energy of the interface between the topmost completed layer and the growing film α_{int} increases so that condition 3.10 eventually breaks down. In this situation up to a certain thickness the film grows in the layer-by-layer manner, then 3D islands start to form on the film surface.

3.3 Fundamental growth processes

Material exchange between the ambient phase and the crystal involves a number of distinct atomic-scale processes. They are generally divided into two groups: the processes in the bulk phases and the processes at the surface


(a) (b) (c)

3.2 Basic modes of film growth: (a) Volmer–Weber mode; (b) Frank–van-der-Merwe mode; (c) Stranski–Krastanov mode.

of the growing crystal. The surface processes control the morphology of the growing film and are therefore worth thorough consideration.

3.3.1 Adsorption and desorption

The very first event in a sequence of elementary surface processes is adsorption of the growing species at the surface. In the adsorption process, the excess energy of the arriving atoms or molecules is accommodated by the crystal, and bonding between the adsorbing particles and the surface is established.

Generally, not every particle that strikes the surface will be adsorbed. Some atoms or molecules arriving at the surface are reflected if their energy cannot be efficiently accommodated by the crystal. In that case only a portion of the incoming flux F of atoms will actually be adsorbed. Introducing the sticking coefficient s, which characterizes the adsorption probability of impinging particles, one can write down the adsorption flux, J_{ads}, as

$$J_{\text{ads}} = sF \tag{3.11}$$

The sticking coefficient is a material-specific property. It depends both on the chemical nature of the adsorbing species and on the local chemistry and structure of the surface. It may also be affected by a number of other factors including the surface temperature, the kinetic energy of the impinging particles, and the adsorbant surface coverage.

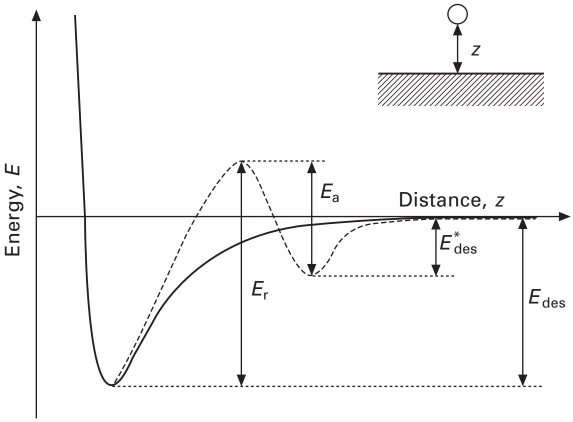
In certain cases adsorption is accompanied by chemical reactions. For instance, adsorption of molecules might involve their dissociation on the surface. In that case adsorption proceeds in several stages. First, a molecule is adsorbed into a weakly bonded precursor state, and then it dissociates into fragments that form stronger bonds to the surface. Adsorption via a precursor state is a thermally activated process characterized by an activation barrier E_a which must be surmounted to go from the precursor state to the final adsorption state (Fig. 3.3).

Adsorbed chemicals may leave the surface by the desorption process. Desorption takes place whenever an atom or molecule gains enough thermal energy to break the bonds with the surface, i.e. desorption is a thermally activated process. The desorption rate of an adsorbed particle is $1/\tau_{\text{des}}$, where τ_{des} is the mean residence time of an adsorbed particle on the surface:

$$1/\tau_{\text{des}} = \nu_0 \exp\left(-\frac{E_{\text{des}}}{k_B T}\right) \tag{3.12}$$

Here $\nu_0 \sim 10^{13}\,\text{s}^{-1}$ is the attempt frequency, E_{des} the activation energy for desorption (Fig. 3.3), and T the substrate temperature.

When particles desorb from the surface in the same chemical form as in



3.3 Energetics of the adsorbtion–desorption processes. The thick solid line represents the potential energy of a particle adsorbing on the surface. The thin dashed line represents the same for the case of thermally activated adsorption via an intermediate precursor state.

the adsorbed state, the desorption reaction obeys first-order kinetics with the desorption rate per unit area (desorption flux) given by

$$J_{des} = \frac{n}{\tau_{des}} \qquad 3.13$$

Here n is the surface concentration of the desorbing particles. First-order desorption is characteristic for simple atomic desorption and desorption of non-dissociated molecules.

Desorption of dissociated molecules is frequently characterized by a higher-order kinetics because the desorption process in that case involves a surface chemical reaction of forming the desorbing molecules from adsorbed atoms. In the steady state the reaction rate per unit area is equal to the desorption flux, leading to

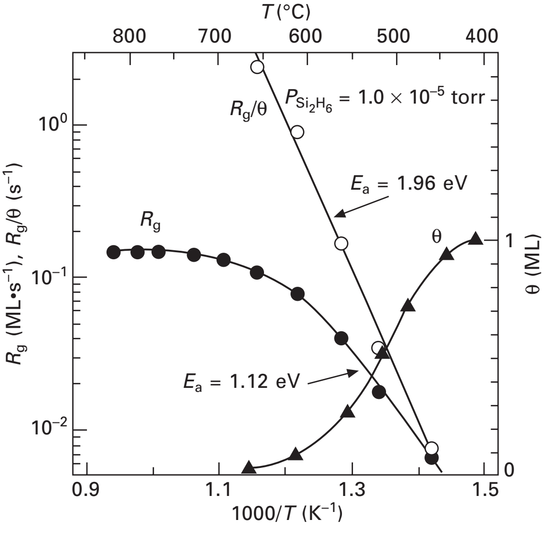
$$J_{des} = k_{des}\, n^r \qquad 3.14$$

Here, k_{des} and r are the rate constant and the order of the reaction, respectively, and n is the concentration of the adsorbed atoms (adatoms). Although k_{des} and r are actually related to the reaction of association of desorbing species from adatoms, they are commonly called the rate constant of desorption and the order of desorption, respectively. Accordingly, the relevant activation barrier, E_r, is commonly referred to as the activation barrier for desorption (Fig. 3.3). For first-order desorption ($r = 1$) Eq. 3.14 is reduced to Eq. 3.13 with $k_{des} = 1/\tau_{des}$.

One typical example of the interplay between adsorption and desorption during growth can be found in the gas-source molecular beam epitaxy

(GSMBE) of silicon using silicon hydrides as the precursor. Since dissociative adsorption of silicon hydrides results in hydrogen termination of surface dangling bonds, growth in GSMBE involves at least the following two steps: (1) adsorption of precursor molecules onto the surface using the dangling bonds, and (2) hydrogen desorption from the surface to regenerate surface dangling bonds. Since hydrogen desorption is thermally activated with an activation energy E_{des} of ~2 eV while molecular adsorption is quite weakly temperature-dependent, the rate-determining step is sharply switched as we increase the growth temperature. Figure 3.4 shows a typical Arrhenius plot of the growth rate obtained by GSMBE using disilane (Nakazawa *et al.*, 1997). The presence of two growth regions, an essentially temperature-independent high-temperature (HT) region and a low-temperature (LT) region with a strong *T*-dependence separated at ~600°C, is evident. The LT region is rate-determined by hydrogen desorption while the HT region is rate-determined by precursor adsorption, which is confirmed by a steep development of the hydrogen coverage as the temperature is decreased below the critical temperature.

The rate determination by hydrogen desorption in LT, however, does not necessarily mean the identity of the growth-rate activation energy E_{gr} in LT with that of hydrogen desorption. This is already evident from the value E_{gr} = 1.12 eV, which is significantly smaller than the reported activation energy for hydrogen desorption from Si surfaces (Narita *et al.*, 2009). A balance between adsorption and desorption of the surface hydrogen yields



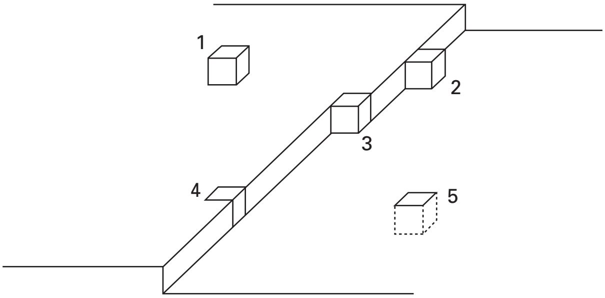
3.4 Temperature dependence of the growth rate of a Si film in GSMBE using disilane.

$$f R_g = \nu \exp\left(-\frac{E_{des}}{k_B T}\right)\theta^r \qquad 3.15$$

where R_g is the growth rate of silicon, f is the number ratio of hydrogen and silicon atoms within a precursor molecule (3 for disilane and 4 for monosilane), and θ is the hydrogen coverage. This simple equation indicates that the activation energy E_{des} for hydrogen desorption should be extracted, not directly from the Arrhenius plot of the growth rate, but from that of R_g/θ, which is shown by open circles in Fig. 3.4. The modified Arrhenius plot shows a good linearity for a wider temperature range, supporting the idea based on Eq. 3.15. More importantly, however, the slope of the line gives an activation energy of 1.96 eV, which is reasonably large for the hydrogen desorption from Si surfaces. A similar story applies also to Si GSMBE using silane, where E_{gr} = 1.29 eV measured from the $R_g(T)$ dependence becomes 1.97 eV in the modified R_g/θ Arrhenius plot, which is in excellent agreement with the value obtained in GSMBE with disilane (Nakazawa *et al.*, 1997).

3.3.2 Surface structure and fundamental growth modes

When an atom or molecule arrives at a crystal surface, it is trapped by a periodic potential imposed by the crystal. The minima of that potential form a regular network of adsorption sites on the surface. There are different types of adsorption sites corresponding to different local structures of the surface and, hence, different coordinations of adsorbed particles. On the atomic scale the structure of a crystalline surface may be described in terms of low-index terraces, steps, and kinks. Accordingly, one may distinguish adsorption sites (1) on the terrace, (2) at the step edge, (3) in the kink, and some other positions with a higher coordination, e.g. (4) in the step edge and (5) in the surface layer. These positions are shown schematically in Fig. 3.5.



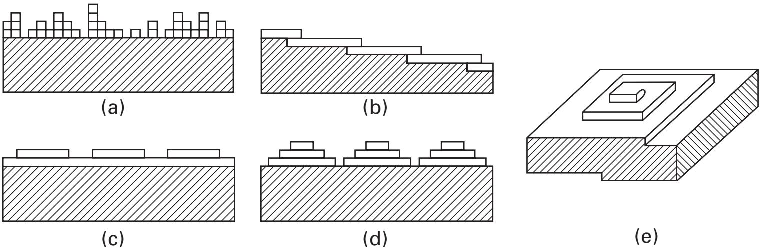
3.5 Different types of adsorption positions on the surface: (1) on a terrace; (2) at the step edge; (3) in the kink; (4) in the step edge; (5) in the surface layer.

Among all the different adsorption sites the kinks play a very special role in crystal growth. Attachment of an atom to the kink site or its removal from the kink does not change the number of broken bonds at the surface and, hence, does not change the free surface energy of the crystal. Therefore the chemical potential of an atom in the kink site is equal to that of the crystal. In other words, the kink is the very place where an adatom becomes a part of the crystal. Along with thermodynamic constraints considered in the previous section, the kink density is one of the major factors that determine the growth kinetics.

If the surface is atomically rough, due to either the crystallography or the entropy effect at elevated growth temperatures, the density of kinks is very high and they could be found at almost any place on the surface. In this case, the adsorbing particles may directly attach to the kink sites, giving rise to the normal mechanism of growth (Fig. 3.6(a)).

On the contrary, if the surface is atomically smooth, the majority of the adsorption sites are terrace sites and the kinks could be found only at the step edges. In this case the crystal grows by lateral motion of the steps, which is generally described as the layer mechanism of growth. Depending on the origin of steps on the surface, one can distinguish four different modes of layer growth, namely step-flow growth, layer-by-layer growth, multilayer or three-dimensional (3D) island growth, and spiral growth. These growth modes are shown schematically in Figs 3.6(b–e).

Step-flow growth is observed on vicinal surfaces. Such surfaces are slightly inclined with respect to the singular low-index crystalline faces, and therefore have a geometry shown in Fig. 3.6(b) with vicinal steps separated by atomically flat low-index terraces. Incorporation of adatoms at the edges of vicinal steps leads to their propagation along the surface. When the kink density provided by vicinal steps is not high enough to swallow up atoms arriving at the surface, additional steps might be created through the nucleation of two-dimensional (2D) islands on the terraces between the vicinal steps.



3.6 Modes of epitaxial growth: (a) normal growth; (b) step-flow growth; (c) layer-by-layer growth; (d) multilayer growth; (e) spiral growth.

This results in a mixed growth mode and eventually to the transition from step-flow to layer-by-layer growth.

In layer-by-layer growth, formation of a new crystalline layer starts with the nucleation of 2D islands on the surface. The islands then grow laterally until they coalesce and cover the whole surface. When a layer is completed, the next layer starts to grow and the whole cycle repeats again (Fig. 3.6(c)).

In certain cases the cyclicity of the layer-by-layer growth breaks up and the growth in the upper layers starts before the lower layers are finished. When this happens due to kinetic constraints, the layer-by-layer growth is replaced by multilayer growth, which appears as the growth of characteristic 3D mounds on the surface (Fig. 3.6(d)). As has been already mentioned in the previous section, thermodynamic constraints may make the growth of 3D islands energetically preferable as the growth scenario, leading to either Volmer–Weber or Stranski–Krastanov growth (Figs 3.2(a) and (c) respectively).

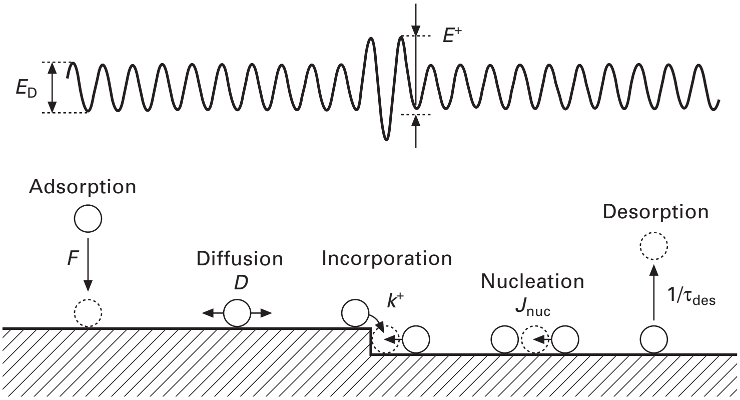
The spiral growth is observed on surfaces with screw dislocations. Each screw dislocation generates a nonvanishing step, which develops a characteristic spiral shape during growth as sketched in Fig. 3.6(e). Since the growth spirals do not disappear during growth, the film may grow without island nucleation.

3.3.3 Surface diffusion and step motion

The motion of surface steps in crystal growth can be described in terms of the Barton, Cabrera, and Frank (BCF) model (Burton *et al.*, 1951), which considers adatoms as a two-dimensional gas and the step edges as sinks for adatoms. Fundamental atomic processes considered in this model are shown in Fig. 3.7. Atoms arrive at the substrate surface with a deposition flux F, diffuse there with a temperature-dependent diffusion coefficient D, and possibly desorb back to the ambient at a rate $1/\tau_{\mathrm{des}}$. If a diffusing atom encounters a step edge it may be incorporated into the crystal if it finds a kink at the step edge. Alternatively, an atom may also become a part of the crystal by taking part in the nucleation of a new 2D or 3D island.

Diffusion on semiconductor surfaces typically occurs by the hopping mechanism, which involves thermally activated diffusion jumps of adatoms from one adsorption site to another with the hopping rate (the jump frequency) $\nu = \nu_0 \exp(-E_{\mathrm{D}}/k_{\mathrm{B}}T)$, where ν_0 is the attempt frequency and E_{D} the surface diffusion barrier (Fig. 3.7). Provided that the concentration of adatoms, n, is small, the macroscopic coefficient of surface diffusion D can be expressed via the microscopic hopping rate ν as $D = a^2\nu/4$, i.e.

$$D = \frac{1}{4} a^2 \nu_0 \exp\left(-\frac{E_{\mathrm{D}}}{k_{\mathrm{B}}T}\right) \tag{3.16}$$



3.7 A one-dimensional representation of the potential energy of an adatom on the surface (top) and atomic-scale surface processes (bottom).

Here a is the jump length (the distance between the adjacent adsorption sites).

In the BCF model the adatom concentration $n(x, y, t)$ satisfies the continuity equation

$$\frac{\partial n}{\partial t} = D\nabla^2 n + F - \frac{n}{\tau_{\text{des}}} \tag{3.17}$$

which accounts for the changes of the adatom population due to the surface diffusion, adsorption and desorption. In writing Eq. 3.17 we assumed for simplicity that the sticking coefficient is unity and the desorption of adatoms from the surface obeys the first-order reaction kinetics. Generalization of the surface diffusion equation to more complicated adsorption/desorption scenarios is straightforward.

This surface diffusion equation (3.17) can be further simplified by noting that the step motion during growth is typically so slow that for each particular step configuration adatoms have plenty of time to establish a quasi-stationary distribution. This justifies the frequently used quasi-steady-state approximation and leads to the steady-state diffusion equation:

$$D\nabla^2 n + F - \frac{n}{\tau_{\text{des}}} = 0 \tag{3.18}$$

The surface diffusion equation should be supplemented by appropriate boundary conditions at the step edges. As a boundary condition at a given step one may naturally use the mass conservation, which implies that there should be a balance between the (1) flux of adatoms entering the step edge from a given terrace by surface diffusion and (2) the net fluxes of adatoms incorporating into the crystal at the given step edge and leaving the terrace by step crossing.

For instance, let us consider a train of straight parallel steps running downhill in the positive direction of the x-axis. For an ith step, situated at x_i, one has (Ozdemir and Zangwill, 1992)

$$\begin{aligned} D\left.\frac{\mathrm{d}n_1}{\mathrm{d}x}\right|_{x=x_i} &= \beta_{\mathrm{inc}}\,[n_1(x_i) - \tilde{n}] + \beta_{\mathrm{p}}\,[n_1(x_i) - n_{\mathrm{u}}(x_i)] \\ -D\left.\frac{\mathrm{d}n_{\mathrm{u}}}{\mathrm{d}x}\right|_{x=x_i} &= \beta_{\mathrm{inc}}\,[n_{\mathrm{u}}(x_i) - \tilde{n}] - \beta_{\mathrm{p}}\,[n_1(x_i) - n_{\mathrm{u}}(x_i)] \end{aligned} \qquad 3.19$$

where $n_1(x)$ and $n_{\mathrm{u}}(x)$ are the adatom concentrations on the lower and upper terraces, respectively, and $\tilde{n}$ is the equilibrium concentration of adatoms on the surface. The first terms on the right-hand sides of Eqs 3.19 represent the net incorporation fluxes, and the second terms are the currents of adatoms that cross the step without visiting the kink. Accordingly, the kinetic coefficients β_{inc} and β_{p} are the incorporation coefficient and the step permeability (or step transparency) coefficient, respectively. The same incorporation coefficient for upper and lower terraces used in Eqs 3.19 presumes a symmetric potential relief around the step edge, as sketched in Fig. 3.7. For a more general case the reader is referred to Filimonov and Hervieu (2004).

By solving Eq. 3.18 with boundary conditions 3.19, one finds the adatom distribution on the terraces between the step edges. Then the step velocity is calculated as

$$v_i = \beta_{\mathrm{inc}}\,[n_1(x_i) - \tilde{n}] + \beta_{\mathrm{inc}}\,[n_{\mathrm{u}}(x_i) - \tilde{n}] \qquad 3.20$$

given the incorporation coefficient β_{inc} and equilibrium adatom concentration $\tilde{n}$.

The common practice regarding the incorporation and the step permeability coefficients is to treat them as phenomenological parameters. This might be a source of confusion if β_{inc} and β_{p} are treated as independent quantities, which is, however, not the case (Pierre-Louis, 2001). To show the mutual dependence of the two parameters, let us consider possible outcomes of an adatom's attachment to a step edge.

Generally, when a diffusing adatom attaches to a step edge, it will be incorporated into the crystal only if it directly attaches to a kink, or if it finds a kink via subsequent migration along the step edge. Therefore the incorporation probability can be written as

$$p_{\mathrm{inc}} = \theta_{\mathrm{k}} + (1 - \theta_{\mathrm{k}})\,f(a) \qquad 3.21$$

where θ_{k} is the density of kinks at the step edge, which determines the probability of the direct attachment to the kink site, and $f(q)$ is the probability that an adatom attached to the step edge at a kink-free step edge position will encounter a kink before detachment from the step edge. The latter probability was estimated in Filimonov and Hervieu (2004) as $f(q) \approx \tanh(q)/q$ where q is the ratio of the average distance between the kinks at the step edge to the mean length of adatom migration along the edge before detachment.

An atom attached to the step edge will detach from the edge with probability $1 - p_{\text{inc}}$ and go to either the lower or the upper terrace. Accordingly, the probability that an adatom will cross the step without visiting the kink is given by

$$p_{\text{cross}} = \tfrac{1}{2}(1 - p_{\text{inc}}) = \tfrac{1}{2}(1 - \theta_{\text{k}})(1 - f(q)) \qquad 3.22$$

where the factor of $\frac{1}{2}$ accounts for two possible detachment directions, whose probabilities are set equal here.

Using Eqs 3.21 and 3.22 one gets the following expressions for the incorporation and permeability coefficients:

$$\beta_{\text{inc}} = p_{\text{inc}}\, a^2 k^+ = [\theta_{\text{k}} + (1 - \theta_{\text{k}}) f(q)] a^2 k^+$$

$$\beta_{\text{p}} = p_{\text{cross}}\, a^2\, k^+ = \tfrac{1}{2}(1 - \theta_{\text{k}})[1 - f(q)] a^2 k^+ \qquad 3.23$$

were $k^+ = \nu_0 \exp(-E^+/k_B T)$ is the rate constant of adatom attachment to the step edge, and E^+ is the activation barrier for adatom attachment (Fig. 3.7).

It is immediately seen from Eqs 3.23 that the incorporation and permeability coefficients cannot be treated as independent quantities. Indeed, in the simple case of symmetric attachment barriers considered here, the following relation holds:

$$\beta_{\text{inc}} + 2\beta_{\text{p}} = a^2 k^+ \qquad 3.24$$

This relation simply reflects the normalization condition on the probabilities for an adatom attaching to the step edge to incorporate into the crystal (p_{inc}) or to detach from the step edge with equal probability (p_{cross}) to either the lower or the upper terrace.

3.3.4 Nucleation

When several diffusing adatoms meet, they form a cluster on the surface. Formation of the cluster costs some energy because a new surface (in the case of 3D clusters) or a new step (in the case of 2D clusters) is created. Small clusters are thermodynamically unstable and rapidly decay because the contributions of the surface and step energies are too large. However, at larger sizes the volume contribution dominates; therefore beyond a certain critical size i^*, the clusters become stable and grow spontaneously by adatom capture. In the thermodynamic nucleation theory the critical cluster size i^* is defined as the size at which the Gibbs free energy of the cluster $\Delta G(i)$

is maximal. The cluster of size i^* is referred to as the critical nucleus. The equilibrium surface concentration of the critical nuclei is given by

$$N^* = n \exp\left(-\frac{\Delta G^*}{k_B T}\right) \tag{3.25}$$

where $\Delta G^* \equiv \Delta G(i^*)$ is the thermodynamic work to create the critical nucleus (see Section 3.2)

At far from equilibrium growth conditions, the critical nucleus might consist of a few atoms. In this case the macroscopic quantities, such as the Gibbs free energy, surface free energy, etc., cannot be defined, and should be replaced by appropriate atomistic terms. For instance, aggregation of unstable clusters from adatoms might be considered as chemical reactions of the form

$$A_1 + A_1 + \ldots \leftrightarrow A_i \tag{3.26}$$

where A_1 denotes the adatoms and A_i a cluster of i atoms. Applying the mass action law to the reaction of formation of the clusters of size i^*, one comes to the so-called Walton relation for the density of i^*-sized unstable clusters:

$$N^* = N_0 \left(\frac{n}{N_0}\right)^{i^*} \exp\left(-\frac{E_{i^*} - i^* E_1}{k_B T}\right) \tag{3.27}$$

Here $N_0 = a^{-2}$ is the density of adsorption sites on the surface, and E_1 and E_{i*} are the potential energies of a single adatom and a cluster of size i^*, respectively. The difference $E_{i*} - i^* E_1$ is the energy required to dissociate the critical nucleus into single adatoms. The Walton relation (Walton, 1962) represents an atomistic analog of Eq. 3.25.

The critical nucleus becomes a stable 2D or 3D island when it captures an additional adatom. Therefore the island nucleation rate J_{nuc} is put proportional to N^* and to the frequency $\omega^* \sim Dn$ of adatom attachment to the critical nuclei:

$$J_{\text{nuc}} = DnN^* \tag{3.28}$$

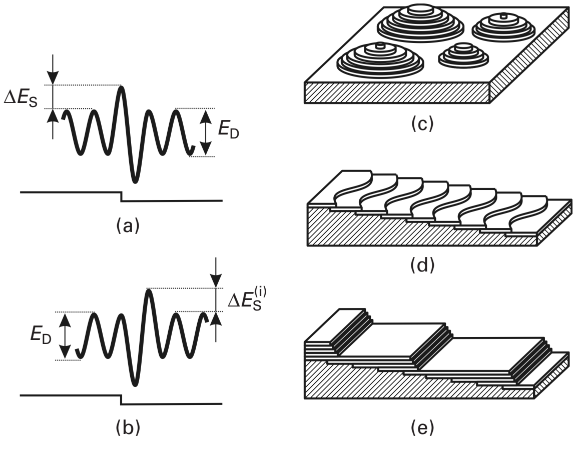
Since the island nucleation rate depends strongly on the adatom concentration n, Eq. 3.28 has to be coupled with an appropriate equation for adatom concentration, e.g. the surface diffusion equation (3.18), to find the island density by integration.

3.4 Kinetics of epitaxial growth

Kinetics of the atomic-scale surface processes go into the forefront in the case of epitaxial growth at far from equilibrium when the growth is too fast and adatoms have no time to find the thermodynamically most favorable places for incorporation. In this case the equilibrium surface morphology cannot be maintained and the growth proceeds in a kinetic growth mode.

A particularly interesting situation arises when kinetic constraints make certain lattice sites less accessible for migrating adatoms. For instance, the downward diffusion of an adatom across the step edge involves a poorly coordinated activated state, therefore attachment of adatoms approaching the step edge from the upper terrace might be hindered by the so-called Ehrlich–Schwoebel (ES) barrier (Ehrlich and Huda, 1966; Schwoebel and Shipsey, 1966). The ES barrier is defined as an extra barrier ΔE_S (i.e. additional to the diffusion barrier E_D) which has to be surmounted by an adatom attaching to the descending step (Fig. 3.8(a)). This barrier has a striking impact on the growth morphology and in certain cases may turn the growth to a kinetically unstable mode.

As a first example let us consider a compact 2D island growing on a high-symmetry (singular) surface. If the downward diffusion of adatoms deposited on top of the island is hindered by the ES barrier, adatoms will accumulate on the island top, thereby increasing the chances of early nucleation of islands in higher layers. This makes the planar layer-by-layer growth kinetically unstable and eventually leads to the formation of multilevel mounds of characteristic 'wedding-cake' shape (Fig. 3.8(c)). This phenomenon is commonly referred to as the Ehrlich–Schwoebel instability (Villain, 1991). When the ES instability is accompanied by the faceting of the mound side walls, the pyramid-like 3D islands form on the surface.



3.8 (a)–(b) Step edge potential barriers: (a) Ehrlich–Schwoebel barrier; (b) inverse Ehrlich–Schwoebel barrier. (c)–(e) Kinetic growth instabilities due to the step-edge barriers: (c) mound formation; (d) step meandering; (e) step bunching.

Looking for experimental evidence for the ES instability, one may refer to numerous observations of the 3D mound formation in homoepitaxial growth of metals (see, for example, a recent review by Evans *et al.* (2006) and references therein). In these systems the mounds appear from the very beginning of the growth, indicating a relatively strong ES barrier, in accordance with results of theoretical calculations. In semiconductor homoepitaxy the 3D mounds were observed on GaAs(001) (Johnson *et al.*, 1994), GaSb(001) (Nosho *et al.*, 2002), and Ge(001) (Van Nostrand *et al.*, 1998). In the latter case, however, the onset of mound formation was delayed to very high film thicknesses, suggesting a weak ES barrier and, probably, an alternative mechanism of mound formation due to the self-crossing of strongly meandering steps (Rost *et al.*, 1996).

On vicinal surfaces growing in the step-flow mode the ES barrier can act as both a stabilizing and a destabilizing factor. To illustrate the stabilizing effect of the ES barrier let us consider a train of vicinal steps where one terrace is somewhat wider than the others. If the downward adatom transport across the step edge is hindered by the ES barrier, the steps are fed mainly by adatoms attaching from the lower terraces. Therefore, the ascending step bordering the wide terrace will get more adatoms than other steps and move faster, thus reducing the width of the leading (wide) terrace. This provides a negative feedback which damps eventual fluctuations of the terrace size and favors growth of equally spaced steps.

While the incorporation asymmetry induced by the ES barrier stabilizes the step flow growth against terrace size fluctuations, it makes the steps kinetically unstable in respect to the shape perturbations (Fig. 3.8(d)). This step instability is commonly referred to as Bales–Zangwill (BZ) instability (Bales and Zangwill, 1990). The physical origin of the BZ instability can be understood as follows. When a small protrusion forms at the step edge it penetrates into the area of a higher adatom concentration and therefore captures more adatoms and grows even faster. On the contrary, a concave portion of the step gains fewer adatoms and slows down. As a result the perturbation is amplified and the step develops a wavy shape. Combined with the stabilizing action of the ES barrier, the BZ instability leads to an in-phase step meandering, which allows the deformed steps to keep the terrace width as uniform as possible.

In-phase meandering of vicinal steps was observed during step-flow growth of Cu on Cu(1,1,17) and Cu(0,2,24) surfaces (Schwenger *et al.*, 1997; Maroutian *et al.*, 2001). However, it was later noted that the temperature and flux dependences of the meander wavelength on these surfaces do not agree with predictions of the Bales–Zangwill theory (Bales and Zangwill, 1990). The idea of an alternative mechanism of step meandering was borrowed from the mounding instability discussed above. The main ingredients of this alternative mechanism are fast diffusion of adatoms along the step edge and a one-dimensional analog of the ES barrier – an additional barrier suppressing diffusion around kinks and outer step corners (Pierre-Louis *et al.*, 1999; Murty and Cooper, 1999). Step meandering due to suppressed corner rounding explains the behavior of the meander wavelength on Cu(1,1,17); however, for Cu(0,2,24) an adequate microscopic mechanism of the observed step meandering is still missing (Kallunki *et al.*, 2002).

On semiconductor surfaces step meandering was reported in Ge(001) (Van Nostrand *et al.*, 1998), GaAs(110) (Tejedor *et al.*, 1998), and Si(111) (Hibino *et al.*, 2003) homoepitaxy. While the observed growth morphology is compatible with the BZ instability, the actual instability mechanism might be affected by the presence of surface reconstructions. An excellent example in that respect is step meandering during growth of Si on Si(111), which is observed only in a very narrow temperature range around the transition temperature T_c between the (1 × 1) and (7 × 7) surface reconstructions (Hibino *et al.*, 2003). The physical origin of this instability is related to the fact that during the (1 × 1) to (7 × 7) phase transition, the (7 × 7) reconstruction preferentially nucleates at the upper step edges (Osakabe *et al.*, 1981). Since adatom diffusion is faster on (1 × 1) than on (7 × 7) (Hibino *et al.*, 2001), adatoms adsorbed at the (1 × 1) portion of the terrace are repelled by a higher diffusion barrier at the boundary between the (1 × 1) and (7 × 7) domains. A narrow stripe of the (7 × 7) structure at the descending step serves as an effective ES barrier that creates a diffusion bias toward the ascending steps and naturally leads to step meandering by the Bales–Zangwill mechanism.

Another type of kinetic growth instability of vicinal surfaces related to the symmetry breaking between the ascending and descending steps bordering a terrace is the kinetic step bunching (Fig. 3.8(e)). During growth the step bunching may occur if adatoms diffusing on terraces preferentially stick to the descending steps. In terms of the step edge barriers this situation is characterized by an inverse ES barrier, $\Delta E_S^{(i)}$, which hinders adatom attachment to the ascending step (see Fig. 3.8(d)). In this case the steps are fed mainly from the upper terraces; therefore wider terraces supply more adatoms to the descending steps and become even wider. Accordingly, the descending steps, terminating narrower terraces, get fewer adatoms and move with even smaller velocities. Eventually, faster steps catch up with slower ones and form step bunches on the surface.

Kinetic step bunching during homoepitaxial growth was observed at certain growth temperatures on vicinal GaAs(110) (Krishnamurthy *et al.*, 1993; Tejedor *et al.*, 1998), Si(111) (Ronda and Berbezier, 2004; Omi *et al.*, 2005), and Si(001) (Schelling *et al.*, 1999) surfaces. While in the first two cases an inverse step edge barrier seems to be the most probable reason for the observed growth instability, the presence of any substantial ES barrier on Si(001) was ruled out by *ab initio* calculations (Zhang *et al.*, 1995). In this latter case the anisotropic kinetics inherent in the (2 × 1) reconstruction of Si(001) induce incorporation asymmetry similar to that induced by the inverse ES barrier. The origin of this asymmetry is easier detachment of adatoms at S_A steps, which in the first place induces pairing of S_A and S_B steps, and then, coupled with the diffusion anisotropy, yields an effective inverse ES barrier (Mysliveček *et al*., 2002).

Apart from the step edge barriers, there are some other factors that may break the symmetry of incorporation fluxes at the ascending and descending steps and cause kinetic instabilities in epitaxial growth. For instance, direct electric current applied to the surface may cause drift of adatoms parallel to the current direction through the electromigration effect (Latyshev *et al*., 1989). Since the drift of adatoms in the step-up direction reduces the incorporation flux into the descending step and increases the flux into the ascending step, it should have a similar effect on the surface morphology as the ES barrier, i.e. it should stabilize vicinal steps against step bunching and cause step meandering. The step-down drift of adatoms favors incorporation into the descending step and, therefore, it might induce kinetic step bunching during growth, for the same reasons as the inverse ES barrier does. Electromigration-induced step bunching has been experimentally observed in Si homoepitaxy on Si(111) surfaces (Métois and Stoyanov, 1999; Gibbons *et al*., 2005).

3.5 Heteroepitaxy

Heteroepitaxy is the growth of a crystalline film on a crystalline substrate of a different material. For most of the practically important combinations of materials, heteroepitaxy is associated with the build-up of elastic strain in the growing film. This strain results from the difference in the lattice constants of the film and the substrate, and influences considerably both the film formation and the electrical properties of the resulting heterostructures.

The quantitative measure of the structural difference between the film and the substrate is the lattice mismatch, defined as the relative difference of the lattice constants of the film (a_f) and the substrate (a_s):

$$f = \frac{a_f - a_s}{a_s} \tag{3.29}$$

For instance, the lattice constant of germanium (0.565 nm) is about 4.2% larger than that of silicon (0.543 nm), which therefore yields a lattice mismatch in the $Si_{1-x}Ge_x$/Si heteroepitaxial system that may vary from zero to 0.042 depending on the Ge concentration x.

The influence of strain on epitaxial growth is twofold. First, strain may affect the growth kinetics by changing the activation barriers for the atomic surface processes. Second, strain may alter thermodynamic properties of the growing film and therefore directly influence the equilibrium growth morphology. The occurrence of Stranski–Krastanow growth of Ge on Si(001) and Si(111) surfaces represents an excellent example in this respect.

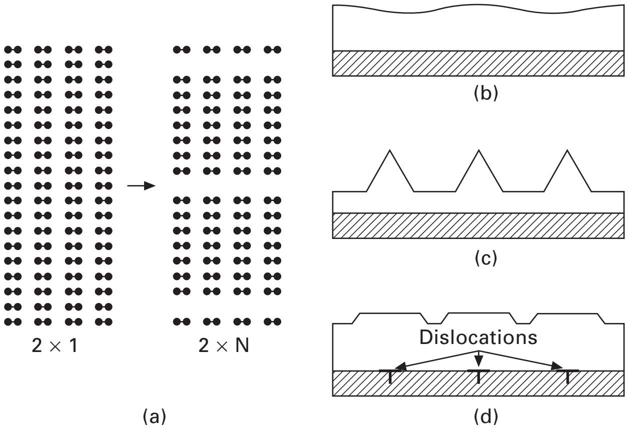
At small film thicknesses and moderate lattice mismatches the film adopts the lattice constant of the substrate. The film is said to be pseudomorphic in this case. During the pseudomorphic growth the film is tetragonally distorted but its surface remains smooth. The lateral positioning of the film's atoms aligns with the substrate at the cost of accumulated strain energy E_{strain}, which increases with the film thickness h as

$$E_{\text{strain}} = \frac{2G(1+\nu)}{1-\nu}\varepsilon^2 h \qquad 3.30$$

Here G is the shear modulus of the film, ν the Poisson ratio, and $\varepsilon = -f$ the misfit strain of the film.

Since the strain energy E_{strain} increases with the film thickness, the pseudomorphic film eventually becomes thermodynamically unstable and should seek a way for strain relaxation at later stages of the growth. At small film thicknesses the increase of strain could be partially accommodated by surface reconstruction. For instance, during growth of a pseudomorphic wetting layer of Ge on Si(001), the (2 × 1) reconstruction of the substrate surface is replaced by a (2 × N) reconstruction in which every Nth dimer of the (2 × 1) structure is missing (Köhler *et al.*, 1992; Iwawaki *et al.*, 1992; Wu *et al.*, 1995). The missing dimers form a regular array of trenches at the surface, which allows outward relaxation of Ge atoms near the trenches, thus reducing the overall strain in the growing film (Tersoff, 1992). The (2 × N) structure first appears with N = 16–17. With increasing surface coverage, however, N gradually decreases, approaching N = 8–9 on a two-monolayer (2 ML) Ge film (Wu *et al.*, 1995; Voigtländer and Kästner, 1999). The decreasing periodicity allows more and more Ge atoms to take relaxed positions, thereby serving to improve strain relaxation. The same arguments seem to apply to the growth of the Ge wetting layer on Si(111), where at certain growth conditions the (7 × 7) reconstruction of the Si(111) surface is replaced by a shorter (5 × 5) structure during growth of a Ge wetting layer on Si(111) (Becker *et al.*, 1985; Köhler *et al.*, 1991; Rosei *et al.*, 2000; Cherepanov and Voigtländer, 2004).

As the strain energy accumulated in the film increases with the increase of the film thickness, a higher degree of strain relief must be sought: elastic deformation of the film. In this case the film develops a corrugated surface morphology (to partially relieve the strain), though it still keeps coherence with the substrate. For instance, a ripple, sketched in Fig. 3.9(b), reduces the strain in the film because atoms at the peaks are less confined and may adopt a larger lattice constant closer to the material's intrinsic one. This might fuel the strain-driven material transfer from more strained (troughs) to less strained (peaks) surface areas, causing further undulation of the grown film.



3.9 Strain relaxation mechanisms in heteroepitaxial growth: (a) change of the surface reconstruction; (b) ripple formation; (c) nucleation of coherent 3D islands; (d) introduction of dislocations.

Although increasing surface energy of the film impedes this tendency, the elastic effect outbalances the surface energy contribution if the corrugation wavelength is longer than a critical wavelength (Asaro and Tiller, 1972; Grinfel'd, 1986; Srolovitz, 1989):

$$\lambda_{\text{crit}} = \frac{1 - \nu}{2G(1 + \nu)^2} \frac{\pi \gamma}{\varepsilon^2} \tag{3.31}$$

This scenario, commonly referred to as the Asaro–Tiller–Grinfeld (ATG) instability, implies gradual development of the strain-relieving surface undulations over an extended period of time. Due to its slow formation kinetics the ATG instability is assumed to be valid only for low-mismatch heteroepitaxial systems, where sufficiently thick dislocation-free heteroepitaxial layers can be grown.

In heteroepitaxial systems characterized by a high lattice mismatch, relaxation of the strain may occur via the nucleation and growth of coherent 3D islands (Eaglesham and Cerullo, 1990) as sketched in Fig. 3.9(c). The 3D island formation causes strain relaxation because atoms at the island side facets are less confined than atoms in the flat wetting layer and therefore the former may take up more comfortable positions. That is, 3D islands reduce strain for the same reason as in the case of surface ripples. The difference, however, is the fact that, in sharp contrast to the ATG instability, the nucleation of 3D islands is a rapid process, which takes place at relatively small film thicknesses. Conversion of a coherent wetting film of volume V into a 3D island of the same volume increases the system free energy by $E_{\text{surf}} \sim \alpha_{\text{S}} V^{2/3}$, whereby the strain energy is reduced by $E_{\text{el}} \sim \varepsilon^2 V$ due to the strain relaxation

(Politi *et al.*, 2000). When the elastic relaxation energy E_{el} outbalances the surface energy E_{surf}, the 3D growth becomes thermodynamically favorable over the planar surface morphology. To reach this thermodynamically preferable state, the system has to overcome a free energy barrier, which scales with the lattice strain as $\sim\varepsilon^4$ (Tersoff and LeGoues, 1994). This dependence implies that the barrier should be relatively low at high mismatches, whereas for less strained systems the 3D nucleation might be inhibited and give way to some other strain relaxation mechanisms, e.g. by introducing dislocations.

When elastic strain relaxation is kinetically suppressed or becomes ineffective, the strain relief proceeds by plastic relaxation (Fig. 3.9(d)). At this stage a set of misfit dislocations forms at the film–substrate interface (for a review on misfit dislocations see Hull and Bean (1992)). The growing film loses registry with the substrate, but instead greatly relaxes its lattice parameter toward the bulk lattice constant of the deposit material. Introduction of the misfit dislocations is the most efficient way to relieve strain, but due to high energy costs associated with the dislocation cores, introduction of dislocations becomes energetically favorable only after a certain critical film thickness h_c. Moreover, since the dislocation formation is a thermally activated process, the onset of plastic relaxation might be delayed to even higher film thickness due to kinetic constraints.

To conclude we will give a brief overview of the growth modes of Ge on Si(111) and Si(001) singular surfaces. More details on SiGe heteroepitaxy and, in particular, on the formation of Ge quantum dots could be found in Chapters 10 and 11.

Growth of Ge on Si(111) starts with the formation of a compressively strained pseudomorphic wetting layer followed by nucleation of 3D islands. The critical thickness for the 2D–3D transition depends on the growth temperature and composition of the depositing film. For MBE growth of pure Ge on Si(111) the critical thickness is about 2–3 bilayers (Denier *et al.*, 1993; Pchelyakov *et al.*, 1997). At higher coverages 3D Ge islands having the shape of truncated pyramids with well-developed facets (Köhler *et al.*, 1991; Voigtländer and Zinner, 1993) nucleate on top of the wetting layer. Initially they keep coherence with the substrate, partially relieving the accumulated strain by outward lattice relaxation at the side facets. For larger islands a transition from coherent to dislocated islands is observed (Motta, 2002). The dislocated islands grow mainly laterally, so that large (111) facets develop on their tops. The presence of the dislocations below such flat islands is indicated by the appearance of characteristic undulations on the island top (Paul *et al.*, 2004; Filimonov *et al.*, 2005). No undulations are seen on the free surface between the islands, which means that the wetting layer remains coherently strained.

In MBE growth of pure Ge on Si(001) the 3D island nucleation starts at a Ge coverage between 3 and 5 monolayers (ML), depending on the Ge flux and substrate temperature. Interestingly, the critical thickness shows a nonmonotonic dependence on the growth temperature, displaying a maximum around 500°C (Nikiforov *et al.*, 2009). In CVD from GeH_4 precursors, a significantly thicker wetting layer (8–9 ML) could be grown (Goldfarb *et al.*, 1997).

The mechanism of the 3D island formation on top of the wetting layer depends crucially on the Ge content of the growing film. If pure Ge is deposited, 3D islands form by a thermally activated nucleation process and from the very beginning appear as elongated 'huts' bounded by (105) planes (Mo *et al.*, 1990). On the contrary, during growth of low-misfit $Si_{1-x}Ge_x$ alloys the 3D islands first appear as small unfaceted mounds or 'prepyramids', which transform to (105) faceted huts or square-based pyramids at larger sizes (Chen *et al.*, 1997; Vailionis *et al.*, 2000; Sutter *et al.*, 2003). As the (105) faceted 3D islands get bigger they change their shape and transform to steeper multifaceted domes, to minimize the free energy of the system (Tomitori *et al.*, 1994; Medeiros-Ribeiro *et al.*, 1998; Ross *et al.*, 1999; Montalenti *et al.*, 2004). In turn, the domes evolve to dislocated super-domes at even larger sizes (Rastelli and von Känel, 2002). In the transient regimes the differently shaped 3D islands may coexist and compete for adatoms available on the surface.

3.6 References

Asaro R J and Tiller W A (1972), 'Interface morphology development during stress corrosion cracking: Part I. Via surface diffusion', *Metall Trans*, 3, 1789–1796.

Bales G S and Zangwill A (1990), 'Morphological instability of a terrace edge during step-flow growth', *Phys Rev B*, 41, 5500–5508.

Becker R S, Golovchenko J A and Swartzentruber B S (1985), 'Tunneling images of the 5 × 5 surface reconstruction on Ge-Si(111)', *Phys Rev B*, 32, 8455–8457.

Burton W K, Cabrera N and Frank F C (1951), 'The growth of crystals and the equilibrium structure of their surfaces', *Phil Trans Roy Soc London, Ser A,* 243, 299–358.

Chen K M, Jesson D E, Pennycook S J, Thundat T and Warmack R J (1997), 'Critical nuclei shapes in the stress-driven 2D-to-3D transition', *Phys Rev B*, 56, R1700-R1703.

Cherepanov V and Voigtländer B (2004), 'Influence of material, surface reconstruction, and strain on diffusion at the Ge(111) surface', *Phys Rev B*, 69, 1–8.

Chernov A A (1984), *Modern Crystallography III. Crystal Growth*, Berlin-Heidelberg-New York, Springer-Verlag.

Denier van der Gon A W, Tromp R M and Reuter M C (1993), 'Low energy electron microscopy studies of Ge and Ag growth on Si(111)', *Thin Solid Films*, 236, 140–145.

Eaglesham D and Cerullo M (1990), 'Dislocation-free Stranski–Krastanow growth of Ge on Si(100)', *Phys Rev Lett*, 64, 1943–1946.

Ehrlich G and Huda F G (1966), 'Atomic view of surface self-diffusion: tungsten on tungsten', *J Chem Phys*, 44, 1039.

Evans J, Thiel P and Bartelt M (2006), 'Morphological evolution during epitaxial thin film growth: Formation of 2D islands and 3D mounds', *Surf Sci Rep*, 61, 1–128.

Filimonov S N and Hervieu Yu Yu (2004), 'Terrace-edge-kink model of atomic processes at the permeable steps', *Surf Sci*, 553, 133–144.

Filimonov S N, Cherepanov V, Paul N, Asaoka H, Brona J and Voigtländer B (2005), 'Dislocation networks in conventional and surfactant-mediated Ge/Si(111) epitaxy', *Surf Sci*, 599, 76–84.

Gibbons B J, Noffsinger J and Pelz J P (2005), 'Influence of Si deposition on the electromigration induced step bunching instability on Si(111)', *Surf Sci*, 575, L51-L56.

Goldfarb I, Owen J H G, Hayden P T, Bowler D R, Miki K and Briggs G A D (1997), 'Gas-source growth of group IV semiconductors: III. Nucleation and growth of Ge/Si(001)', *Surf. Sci*, 394, 105–118.

Grinfel'd M Ya (1986), 'Instability of the separation boundary between a nonhydrostatically stressed elastic body and a melt', *Dokl Akad Nauk SSSR*, 290, 1358–1363 [(1986), *Sov Phys Dokl*, 31, 831–834].

Hibino H, Hu C-W, Ogino T and Tsong I (2001), 'Diffusion barrier caused by 1×1 and 7×7 on Si(111) during phase transition', *Phys Rev B*, 64, 245401.

Hibino H, Homma Y, Uwaha M and Ogino T (2003), 'Step wandering induced by homoepitaxy on Si(111) during "1×1"–7×7 phase transition', *Surf Sci*, 527, L222–L228.

Hull R and Bean J (1992), 'Misfit dislocations in lattice-mismatched epitaxial films', *Crit Rev Solid State Mater Sci*, 17, 507–546.

Iwawaki F, Tomitori M and Nishikawa O (1992), 'STM study of initial stage of Ge epitaxy on Si(001)', *Ultramicroscopy*, 42–44, 902–909.

Johnson M D, Orme C, Hunt A W, Graff D, Sudijono J, Sander L M and Orr B G (1994), 'Stable and unstable growth in molecular beam epitaxy', *Phys Rev Lett*, 72, 116–119.

Kallunki J, Krug J and Kotrla M (2002), 'Competing mechanisms for step meandering in unstable growth', *Phys Rev B*, 65, 1–7.

Köhler U, Jusko O, Pietsch G, Müller B and Henzler M (1991), 'Strained-layer growth and islanding of germanium on Si(111)-(7 × 7) studied with STM', *Surf Sci*, 248, 321–331.

Köhler U, Jusko O, Müller B, Horn-von-Hoegen M and Pook M (1992), 'Layer-by-layer growth of germanium on Si(100): strain-induced morphology and the influence of surfactants', *Ultramicroscopy*, 42–44, 832–837.

Krishnamurthy M, Wassermeier M, Williams D R M and Petroff P M (1993), 'Periodic faceting on vicinal GaAs(110) surfaces during epitaxial growth', *Appl Phys Lett*, 62, 1922–1924.

Latyshev A V, Aseev A L, Krasilnikov A B and Stenin S I (1989), 'Transformations on clean Si(111) stepped surface during sublimation', *Surf Sci*, 213, 157–169.

Markov I V (2003), *Crystal Growth for Beginners: Fundamentals of Nucleation, Crystal Growth and Epitaxy*, 2nd edn, Singapore, World Scientific.

Maroutian T, Douillard L and Ernst H-J (2001), 'Morphological instability of Cu vicinal surfaces during step-flow growth', *Phys Rev B*, 64, 165401-1–165401-7.

Medeiros-Ribeiro G, Bratkovski A M, Kamins T, Ohlberg D A A and Williams R S (1998), 'Shape transition of germanium nanocrystals on a silicon (001) surface from pyramids to domes', *Science*, 279, 353–355.

Métois J and Stoyanov S (1999), 'Impact of the growth on the stability–instability transition at Si(111) during step bunching induced by electromigration', *Surf Sci*, 440, 407–419.

Michely T and Krug J (2003), *Islands, Mounds, and Atoms: Patterns and Processes in Crystal Growth Far from Equilibrium*, Berlin, Springer.

Mo Y-W, Savage D E, Swartzentruber B S and Lagally M G (1990), 'Kinetic pathway in Stranski–Krastanov growth of Ge on Si(001)', *Phys Rev Lett,* 65, 1020-1023.

Montalenti F, Raiteri P, Migas D B, von Känel H, Rastelli A, Manzano C, Costantini G, Denker U, Schmidt O G, Kern K and Miglio L (2004), 'Atomic-scale pathway of the pyramid-to-dome transition during Ge growth on Si(001)', *Phys Rev Lett*, 93, 216102-1–216102-4.

Motta N (2002), 'Self-assembling and ordering of Ge/Si(111) quantum dots: scanning microscopy probe studies', *J Phys Condens Matter*, 14, 8353–8378.

Murty R M V and Cooper B H (1999), 'Instability in molecular beam epitaxy due to fast edge diffusion and corner diffusion barriers', *Phys Rev Lett*, 83, 352–355.

Myslivecček J, Schelling C, Schäffler F, Springholz G, Šmilauer P, Krug J and Voigtländer B (2002), 'On the microscopic origin of the kinetic step bunching instability on vicinal Si(0 0 1)', *Surf Sci*, 520, 193–206.

Nakazawa H, Suemitsu M and Miyamoto N (1997), 'Effects of adsorption kinetics on the low-temperature growth-rate activatiion energy in Si gas-source molecular beam epitaxy', *Jpn J Appl Phys*, 36, L703–L704.

Narita Y, Kihara Y, Inanaga S and Namiki A (2009), 'Substantially low desorption barriers in recombinative desorption of deuterium from a Si(100) surface', *Surf Sci*, 603, 1168–1174, and references therein.

Nikiforov A I, Ulyanov V V, Timofeev V A and Pchelyakov O P (2009), 'Wetting layer formation in superlattices with Ge quantum dots on Si(100)', *Microelectronics J*, 40, 782–784.

Nosho B Z, Bennett B R, Aifer E H and Goldenberg M (2002), 'Surface morphology of homoepitaxial GaSb films grown on flat and vicinal substrates', *J Crystal Growth*, 236, 155–164.

Omi H, Homma Y, Tonchev V and Pimpinelli A (2005), 'New types of unstable step-flow growth on Si(111)-(7 ×7) during molecular beam epitaxy: scaling and universality', *Phys Rev Lett*, 95, 216101-1–216101-4.

Osakabe N, Tanishiro Y, Yagi K and Honjo G (1981), 'Direct observation of the phase transition between the (7 × 7) and (1 × 1) structures of clean (111) silicon surfaces', *Surf Sci*, 109, 353–366.

Ozdemir M and Zangwill A (1992), 'Morphological equilibration of a facetted crystal', *Phys Rev B*, 45, 3718–3729.

Paul N, Asaoka H, Mysliveček J and Voigtländer B (2004), 'Growth mechanisms in Ge/Si(111) heteroepitaxy with and without Bi as a surfactant', *Phys Rev B*, 69, 193402(4).

Pchelyakov O P, Markov V A, Nikiforov A I and Sokolov L V (1997), 'Surface processes and phase diagrams in MBE growth of Si/Ge heterostructures', *Thin Solid Films*, 306, 299–306.

Pierre-Louis O (2001), 'Continuum model for low temperature relaxation of crystal steps', *Phys Rev Lett*, 87, 106104(4).

Pierre-Louis O, D'Orsogna M R and Einstein T L (1999), 'Edge diffusion during growth: the kink Ehrlich–Schwoebel effect and resulting instabilities', *Phys Rev Lett*, 82, 3661–3664.

Pimpinelli A and Villain J (1998), *Physics of Crystal Growth*, Cambridge, Cambridge University Press.

Politi P, Grenet G, Marty A, Ponchet A and Villain J (2000), 'Instabilities in crystal growth by atomic or molecular beams', *Phys Rep*, 324, 271–404.

Rastelli A and von Känel H (2002), 'Surface evolution of faceted islands', *Surf Sci*, 515, L493–L498.

Ronda A and Berbezier I (2004), 'Self-patterned Si surfaces as templates for Ge islands ordering', *Physica E: Low-Dimensional Systems and Nanostructures*, 23, 370–376.

Rosei F, Motta N, Sgarlata A, Capellini G and Boscherini F (2000), 'Formation of the wetting layer in Ge/Si(111) studied by STM and XAFS', *Thin Solid Films*, 369, 29–32.

Ross F M, Tromp R M and Reuter M C (1999), 'Transition states between pyramids and domes during Ge/Si island growth', *Science*, 286, 1931–1934.

Rost M, Šmilauer P and Krug J (1996), 'Unstable epitaxy on vicinal surfaces', *Surf Sci*, 369, 393–402.

Schelling C, Springholz G and Schäffler F (1999), 'Kinetic growth instabilities on vicinal Si(001) surfaces', *Phys Rev Lett*, 83, 995–998.

Schwenger L, Folkerts R L and Ernst H J (1997), 'Bales–Zangwill meandering instability observed in homoepitaxial step-flow growth', *Phys Rev B*, 55, 7406–7409.

Schwoebel R L and Shipsey E J (1966), 'Step motion on crystal surfaces', *J Appl Phys*, 37, 3682–3686.

Srolovitz D J (1989), 'On the stability of surfaces of stressed solids', *Acta Metallurgica*, 37, 621–625.

Sutter P, Zahl P and Sutter E (2003), 'Continuous formation and faceting of SiGe islands on Si(100)', *Appl Phys Lett*, 82, 3454–3456.

Tejedor P, Allegretti F E, Šmilauer P and Joyce B A (1998), 'Temperature-dependent unstable homoepitaxy on vicinal GaAs(110) surfaces', *Surf Sci*, 407, 82–89.

Tersoff J (1992), 'Missing dimers and strain relief in Ge films on Si(100)', *Phys Rev B*, 45, 8833–8836.

Tersoff J and LeGoues F K (1994), 'Competing relaxation mechanisms in strained layers', *Phys Rev Lett*, 72, 3570–3573.

Tomitori M, Watanabe K, Kobayashi M and Nishikawa O (1994), 'STM study of the Ge growth mode on Si(001) substrates', *Appl Surf Sci*, 76–77, 322–328.

Vailionis A, Cho B, Glass G, Desjardins P, Cahill D G and Greene J E (2000), 'Pathway for the strain-driven two-dimensional to three-dimensional transition during growth of Ge on Si(001)', *Phys Rev Lett*, 85, 3672–3675.

Van Nostrand J E, Chey S J and Cahill D G (1998), 'Low-temperature growth morphology of singular and vicinal Ge(001)', *Phys Rev B*, 57, 12536–12543.

Villain J (1991), 'Continuum models of crystal growth from atomic beams with and without desorption', *J Physique I*, 1, 19–42.

Voigtländer B and Kästner M (1999), 'Evolution of the strain relaxation in a Ge layer on Si(001) by reconstruction and intermixing', *Phys Rev B*, 60, R5121-R5124.

Voigtländer B and Zinner A (1993), 'Simultaneous molecular beam epitaxy growth and scanning tunneling microscopy imaging during Ge/Si epitaxy', *Appl Phys Lett*, 63(22), 3055.

Walton D (1962), 'Nucleation of vapor deposits', *J Chem Phys*, 37, 2182–2188.

Wu F, Chen X, Zhang Z and Lagally M G (1995), 'Reversal of step roughness on Ge-covered vicinal Si(001)', *Phys Rev Lett*, 74, 574–577.

Zhang Q M, Roland C, Bogusławski P and Bernholc J (1995), 'Ab initio studies of the diffusion barriers at single-height Si(100) steps', *Phys Rev Lett*, 75, 101–104.

4

Types of silicon–germanium (SiGe) bulk crystal growth methods and their applications

N. USAMI, Tohoku University, Japan

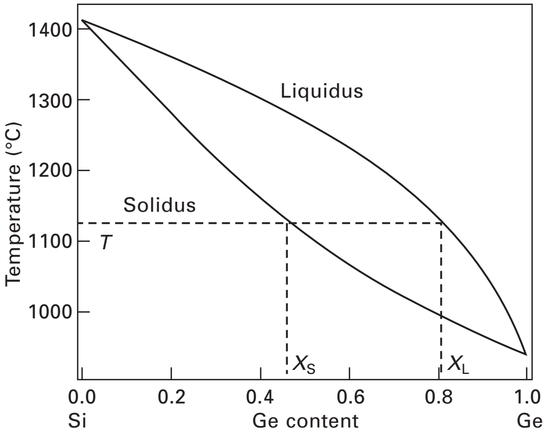
Abstract: This chapter gives a brief overview of the growth of SiGe bulk crystal and its application to substrates for heteroepitaxy. The chapter starts with a description of the requirements for the crystal growth technology to realize single crystalline SiGe bulk crystal with uniform Ge content. It then describes various attempts to control Ge content and avoid polycrystallization. Finally, a couple of examples to utilize SiGe bulk substrates are introduced.

Key words: Czochralski method, float zone method, multicomponent zone melting method, constitutional supercooling, inhomogeneous nucleation, preferential orientation.

4.1 Introduction

As we have already learned in Chapter 1, SiGe is a completely miscible system [1] and therefore SiGe crystals with any Ge contents can be realized in principle. This offers opportunities for scientists and engineers to control material properties and improve device performance by incorporating SiGe to Si. However, crystal growth technology of single crystalline SiGe bulk crystal with uniform Ge content from molten SiGe has not yet been established. This is in contrast to many excellent outcomes using SiGe heterostructures based on vapor phase epitaxial growth of SiGe thin film crystals under non-equilibrium growth conditions. This is due to the intrinsic difficulties of controlling Ge content and avoiding polycrystallization during bulk crystal growth from molten SiGe. This chapter starts by describing requirements for crystal growth technology to realize SiGe bulk crystals with uniform Ge content based on the equilibrium phase diagram shown in Fig. 4.1, and then discusses possible mechanisms of polycrystallization.

The phase diagram of SiGe is very simple and contains only two lines, the liquidus and solidus lines. These two lines intersect at the melting points of Si (1414°C) and Ge (938°C). Generally, melt growth proceeds under near-equilibrium growth conditions, and Ge content in the $Si_{1-x}Ge_x$ crystal, x_S, grown from molten SiGe at a given temperature, T, is determined by the solidus line (see Fig. 4.1). Therefore, we need to control the growth



4.1 Phase diagram of SiGe: the Ge content in the SiGe bulk crystal grown from the melt at a given temperature *T* can be ascertained from the solidus line.

temperature at a constant value to realize uniform Ge content. Due to the large difference between the melting points of Si and Ge, the liquidus and solidus lines are strongly dependent on temperature. As a consequence, a small fluctuation in the temperature, ΔT, could result in a large variation in the Ge content, Δx_S. Furthermore, it is noted that there is a large separation between the liquidus and solidus lines at a given temperature. This means that the segregation coefficient, $k = x_S/x_L$, is not unity, and the solidus Ge content, x_S, is smaller than the liquidus Ge content, x_L. Therefore, if we start with a finite amount of molten SiGe, crystal growth results in depletion of Si from the melt, causing the driving force of the crystal growth to disappear. Therefore, there are two requirements for the crystal growth technology to realize SiGe bulk crystal with uniform Ge content: to keep the growth temperature constant, and to somehow feed Si in the SiGe melt.

In addition, we need to grow single crystalline SiGe instead of polycrystalline SiGe. In contrast to the crystal growth of a single element, there are intrinsic difficulties in realizing a single crystal in the melt growth of plural elements such as SiGe. Owing to the segregation of a component during crystal growth, Ge content in the SiGe melt is not uniform, especially around the solid–liquid interface. This could lead to constitutional supercooling if the temperature gradient is not appropriate, and could result in polycrystallization. Furthermore, due to the difference between the lattice constants of Si and Ge, inhomogeneous built-in strain might be introduced in the grown crystal if the control of the Ge content is not successful. Relaxation of such an internal strain could be a driving force of polycrystallization. These two fundamental difficulties in realizing a single crystal are inherent in the melt growth of plural elements. Of course, careful attention must be paid to other common mechanisms such as inhomogeneous nucleation at the contact with the crucible.

In the following sections, various attempts to grow SiGe bulk crystals from molten SiGe are introduced.

4.2 Growth methods

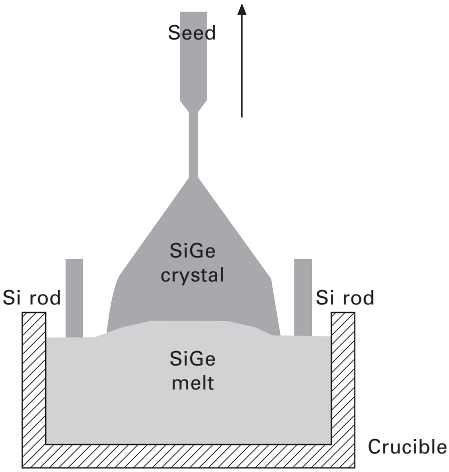
4.2.1 Czochralski method

The Czochralski (CZ) method is a crystal growth technology that starts with insertion of a small seed crystal into a melt in a crucible, pulling the seed upwards to obtain a single crystal. The method is named after the Polish scientist Jan Czochralski, who developed it in 1916. The method is nowadays used for production of various single crystals such as Si and Ge.

Several researchers have applied the CZ method to obtain single crystalline SiGe. Kürten *et al.* carried out the CZ growth of Si-rich SiGe using an extremely slow pulling rate to achieve a stable liquid–solid interface. In addition, the gas flow geometry was designed to achieve good stability of the temperature. As a consequence, single crystalline SiGe with Ge content up to 0.2 was achieved [2]. Yonenaga *et al.* have intensively studied the CZ growth of $Si_{1-x}Ge_x$ in the whole range of Ge content, and obtained single crystals for $0 < x < 0.1$ and $0.85 < x < 1$ [3–8]. For the intermediate Ge content, the grown crystals were polycrystals. They estimated the critical growth velocity of the occurrence of constitutional supercooling as a function of Ge content at a given temperature gradient based on Tiller's criterion [9]. However, even when the growth velocity is slower than the critical value, polycrystallization was observed for the intermediate Ge content. Regarding uniformity of Ge content, no special attention was paid and the depletion of Si was observed along the pulling direction.

A modified CZ method was attempted by Deitch *et al.* [10]. They used pre-grown large-diameter single crystal Si seeds to reduce cap crystallization time and to ensure single crystallinity at a large crystal diameter. In fact, SiGe single crystals up to 68 mm in diameter were realized for Ge content up to 0.17.

In order to avoid depletion of Si during the CZ growth of SiGe, Abrosimov *et al.* have used Si rods for continuous feeding of Si to the melt (Fig. 4.2) [11]. By controlling the movement rate of the Si rods and the pulling rate of the grown crystal, they have attempted to keep the Ge content in SiGe uniform. Global modeling of the whole process, including dissolution of Si through the solid–liquid interface and the transport in the melt, will be necessary to improve the process.



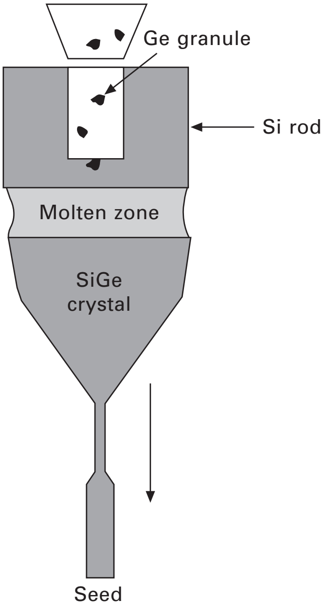
4.2 Schematic of the CZ growth of SiGe bulk crystal with continuous feeding of Si to the melt.

4.2.2 Float zone method

The float zone (FZ) method is a crystal growth technology that creates a localized molten zone where the crystal grows. The starting material is a polycrystalline rod and a seed crystal. Unlike the CZ method, crystal growth is performed without any contact with the crucible wall, which drastically reduces the concentrations of impurities, and high-purity crystals can be obtained. Therefore, FZ Si is generally used for power devices. The diameter of the ingot is limited by the surface tension of the molten zone and is generally not greater than 150 mm.

In terms of crystal growth of SiGe, the advantages of the FZ method are the sharp temperature gradient to avoid consitutional supercooling and the lack of wall contact. Both are useful to suppress polycrystallization. However, it is difficult to feed the base elements in the molten zone.

Several researchers have tried to grow SiGe bulk crystals by the FZ method. Saidov *et al.* used the electron beam FZ technique for pre-synthesized polycrystalline SiGe source and Si(111) seed crystal, and obtained single crystalline SiGe with Ge content up to 0.5. The diameter is limited to less than 3 mm [12]. Wollweber *et al.* used the RF-heated FZ method. As a special technique, axially drilled Si feed rods were employed and Ge granules were continuously fed during the growth (Fig. 4.3). As a consequence, single crystals of SiGe were obtained with Ge content up to 0.22 [13, 14]. Campbell *et al.* have grown Ge-rich SiGe single crystals by the FZ method using a monoellipsoid mirror furnace [15]. They utilized pre-synthesized $Si_{0.05}Ge_{0.95}$ polycrystalline material and a Ge seed. Morphological instabilities

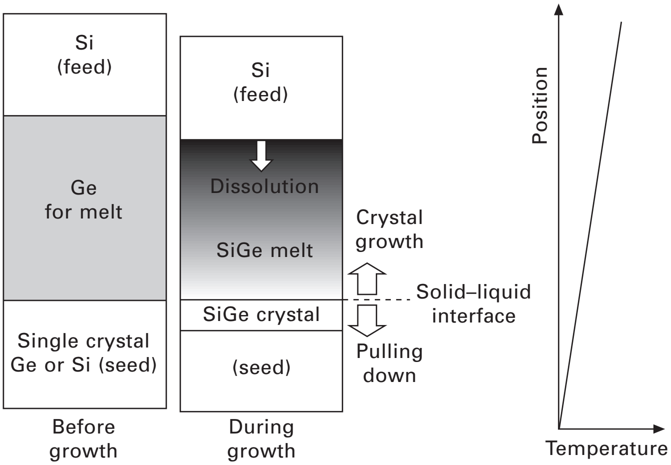


4.3 Schematic of FZ growth of SiGe bulk crystal with continuous charging of Ge granules.

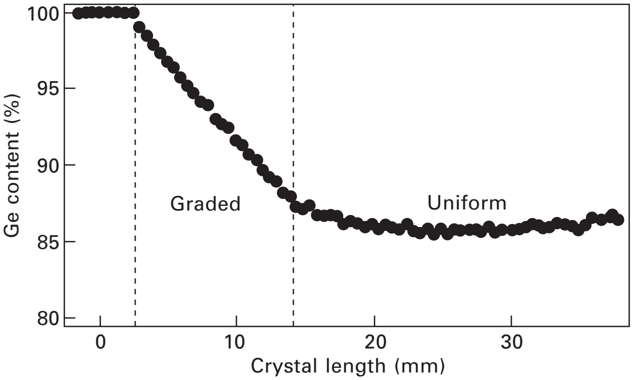
were observed depending on the Si content, and they have speculated that solutal Marangoni convection plays a role.

4.2.3 Multicomponent zone melting method

As a crystal growth method to satisfy the two requirements to realize uniform alloy content in the growth of multicomponent semiconductors, Suzuki, Nakajima *et al*. proposed the 'multicomponent zone melting method' [16]. Figure 4.4 shows the arrangement of raw materials before the growth, the schematic distribution of the material composition during the growth, and the temperature distribution. To start the growth, a seed crystal (Ge or Si), polycrystalline Ge for preparing the melt, and a Si crystal to continuously feed Si to the melt are set in a crucible. By subjecting the crucible to an appropriate temperature gradient, the molten Ge is sandwiched between the seed and the feed crystal. By dissolution of Si to the melt, a binary SiGe melt is formed. Si diffuses toward the solid–liquid interface driven by the concentration gradient, and the resultant supersaturation drives the crystal growth. The depletion of Si is not a concern any more since Si is continuously fed to the melt through dissolution. Furthermore, growth at a constant temperature can be realized by balancing the growth rate and the


Si
(feed)
Ge
for melt
Single crystal
Ge or Si (seed)
Before
growth
Si
(feed)
Dissolution
SiGe melt
SiGe crystal
(seed)
During
growth
Crystal
growth
Solid–liquid
interface
Pulling
down
Position
Temperature

4.4 Schematic of the fundamental concept of the multicomponent zone melting method. The growth temperature is kept constant by balancing the pulling rate of the crucible with the growth rate. Si is continuously fed to the melt through dissolution during crystal growth.

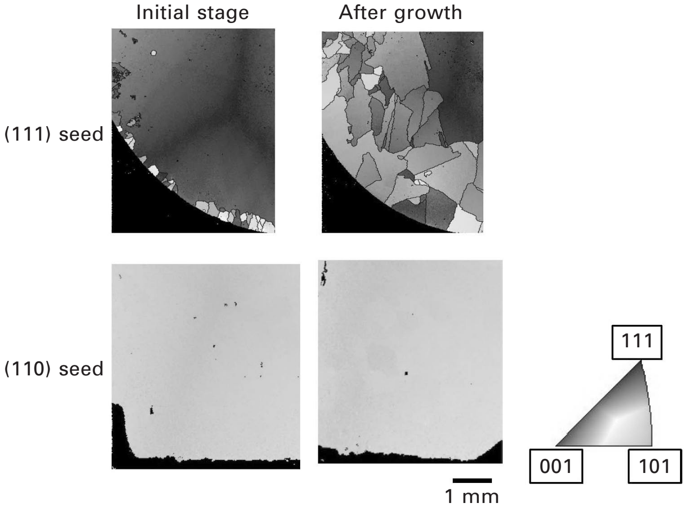


4.5 Typical distribution of the Ge content in the SiGe bulk crystal grown by the multicomponent zone melting method.

rate to pull down the crystal, leading to realization of SiGe bulk crystal with uniform Ge content.

To implement the fundamental concept, integration of a series of improvements has been necessary, such as the *in-situ* monitoring system of the interface position [17–19], the dynamic control of the pulling rate [20], and so on. Figure 4.5 shows a typical distribution of Ge content in the SiGe crystal along the growth direction grown by the multicomponent zone melting method. The initial gradual increase of the Si content reflects the gradual increase of the growth temperature, since growth was achieved without moving the crucible at the initial stage. Then, pulling down of the crucible was started, to balance the growth rate. As a result, Ge content is seen to be successfully controlled along the growth direction. Similar growth methods within the family of 'directional solidification' have also been investigated by other research groups under the names of 'liquid phase diffusion [21]', 'traveling liquidus-zone method [22, 23]', and so on [24–26].

In these methods, polycrystallization has been an issue due to the contact with the crucible wall and/or the internal stress coming from the inhomogeneous distribution of Ge. To suppress the polycrystallization, Azuma *et al.* attempted to optimize crystallographic orientation of the seed to the preferential orientation [27]. In order to determine the preferential orientation, they carried out a simple experiment of directional solidification on multiple seed crystals with random orientations, and investigated the orientation distribution of the grown crystal by electron back-scattering diffraction pattern analysis. Based on the results, they specified that <110> is the preferential orientation of SiGe under the growth conditions. Figure 4.6 shows the impact of the orientation of the seed during growth of SiGe bulk crystals. When non-preferential Ge(111) was used as the seed crystal, small crystal grains originating from the inhomogeneous nucleation at the crucible wall became large as crystal growth proceeded. On the other hand,



4.6 Comparison of orientation distribution of SiGe bulk crystals grown by the multicomponent zone melting method using non-preferential Ge(111) seed (upper two images) and preferential Ge(110) seed (lower two images). It is clearly seen that polycrystallization is suppressed by using the preferential Ge(110) seed.

by using Ge(110) as the seed, polycrystallization was found to be suppressed. As illustrated by these examples, the optimization of the orientation of the seed crystal is of crucial importance to obtain single crystalline SiGe bulk crystals. It should also be noted that the use of $CaCl_2$ as a liquid encapsulant is shown to be useful to avoid inhomogeneous nucleation owing to the contact with the crucible wall [28, 29].

4.3 Application of silicon–germanium (SiGe) bulk crystal to heteroepitaxy

A couple of examples to utilize SiGe bulk substrates for heteroepitaxy are introduced. Usami *et al.* utilized Si-rich SiGe bulk crystal as a substrate for strained Si thin film [30]. Epitaxial growth of Si was carried out both on a SiGe bulk substrate and on a conventional SiGe virtual substrate. By X-ray reciprocal space mapping and analysis of microscopic Raman spectra, fluctuation of the orientation and the strain in the strained Si layer was revealed to be drastically suppressed by exploiting SiGe bulk substrate. They have also grown a couple of strained Si/strained Ge quantum wells on a CZ-grown $Si_{0.915}Ge_{0.085}$ bulk substrate, and observed photoluminescence of excitons confined in the quantum wells without any dislocation-related luminescence [31]. Tsujino *et al.* have fabricated *p*-type SiGe double barrier resonant tunneling diodes on a SiGe substrate grown by the multicomponent zone melting method. Devices exhibited a high peak-to-valley current ratio up to 8.8 at 4.2 K and a negative differential resistance up to 340 K [32]. Nihei *et al.* obtained a SiGe bulk crystal with a graded Ge content by intentionally changing the growth temperature, and utilized it as a substrate with a lateral variation in the Ge content [33]. This allows a systematic change in lattice mismatch between the film and the substrate in a single sample. In fact, they observed that the morphology of the grown thin film crystals is strongly dependent on the amount and the sign of the lattice mismatch.

These examples show that SiGe bulk substrates have a potential for various research and development projects, such as fabrication of high-performance SiGe heterostructure devices based on the quantum effect, fundamental research to pursue crystal growth mechanisms and material properties, and so on.

4.4 Conclusion

A brief overview of the growth of SiGe bulk crystal was given. In order to control the Ge content in the SiGe bulk crystal, we need to keep the growth temperature constant and to somehow feed Si to the melt. Small-scale SiGe bulk crystals with almost uniform Ge content have been realized by various crystal growth techniques to satisfy these two requirements. In addition, we need to avoid polycrystallization during crystal growth. Due to the inherent difficulties of controlling the uniformity of the Ge content and obtaining single crystals, it would be difficult to achieve a large-scale single crystalline SiGe ingot with uniform composition. However, SiGe bulk crystals have great potential to contribute to the progress of both fundamental and applied research related to SiGe heterostructures.

4.5 References

[1] Olesinski, R. W. & Abbaschian, G. J. 1984. *Bull. Alloy Phase Diagram*, 5, 180.

[2] Kürten, M. & Schilz, J. 1994. Czochralski growth of Si_xge_{1-x} single-crystals. *Journal of Crystal Growth*, 139, 1–5.

[3] Yonenaga, I., Matsui, A., Tozawa, S., Sumino, K. & Fukuda, T. 1995. Czochralski Growth of $Ge_{1-x}Si_x$ alloy crystals. *Journal of Crystal Growth*, 154, 275–279.

[4] Matsui, A., Yonenaga, I. & Sumino, M. 1998. Czochralski growth of bulk crystals of $Ge_{1-x}Si_x$ alloys. *Journal of Crystal Growth*, 183, 109–116.

[5] Yonenaga, I. & Nonaka, M. 1998. Czochralski growth of bulk crystals of $Ge_{1-x}Si_x$ alloys II. Si-rich alloys. *Journal of Crystal Growth*, 191, 393–398.

[6] Yonenaga, I. 1999. Czochralski growth of GeSi bulk alloy crystals. *Journal of Crystal Growth*, 198, 404–408.

[7] Yonenaga, I. 2001. Czochralski growth of heavily impurity doped crystals of GeSi alloys. *Journal of Crystal Growth*, 226, 47–51.

[8] Yonenaga, I. 2005. Growth and fundamental properties of SiGe bulk crystals. *Journal of Crystal Growth*, 275, 91–98.

[9] Tiller, W. A., Jackson, K. A., Rutter, J. W. & Chalmers, B. 1953. The Redistribution of solute atoms during the solidification of metals. *Acta Metallurgica*, 1, 428–437.

[10] Deitch, R. H., Jones, S. H. & Digges, T. G. 2000. Bulk single crystal growth of silicon–germanium. *Journal of Electronic Materials*, 29, 1074–1078.

[11] Abrosimov, N. V., Rossolenko, S. N., Thieme, W., Gerhardt, A. & Schröder, W. 1997. Czochralski growth of Si- and Ge-rich SiGe single crystals. *Journal of Crystal Growth*, 174, 182–186.

[12] Saidov, M. S., Yusupov, A. & Umerov, R. S. 1981. Si–Ge solid-solution single-crystal growth by electron-beam floating zone technique. *Journal of Crystal Growth*, 52, 514–518.

[13] Wollweber, J., Schulz, D. & Schröder, W. 1996. Extremely reduced dislocation density in Si_xGe_{1-x} single crystals grown by the float zone technique. *Journal of Crystal Growth*, 158, 166–168.

[14] Wollweber, J., Schulz, D. & Schröder, W. 1996. Si_xGe_{1-x} single crystals grown by the RF-heated float zone technique. *Journal of Crystal Growth*, 163, 243–248.

[15] Campbell, T. A., Schweizer, M., Dold, P., Croll, A. & Benz, K. W. 2001. Float zone growth and characterization of $Ge_{1-x}Si_x$ ($x <= 10$ at%) single crystals. *Journal of Crystal Growth*, 226, 231–239.

[16] Suzuki, T., Nakajima, K., Kusunoki, T. & Katoh, T. 1996. Multicomponent zone melting growth of ternary InGaAs bulk crystal. *Journal of Electronic Materials*, 25, 357–361.

[17] Usami, N., Azuma, Y., Ujihara, T., Sazaki, G., Miyashita, S., Murakami, Y. & Nakajima, K. 2001. Growth of Si_xGe_{1-x} ($x = 0.15$) bulk crystal with uniform composition utilizing in situ monitoring of the crystal–solution interface. *Japanese*

Journal of Applied Physics Part 1 – Regular Papers, Short Notes and Review Papers, 40, 4141–4144.

[18] Azuma, Y., Usami, N., Ujihara, T., Sazaki, G., Murakami, Y., Miyashita, S., Fujiwara, K. & Nakajima, K. 2001. Growth of SiGe bulk crystal with uniform composition by directly controlling the growth temperature at the crystal–melt inter-face using in situ monitoring system. *Journal of Crystal Growth*, 224, 204–211.

[19] Sazaki, G., Azuma, Y., Miyashita, S., Usami, N., Ujihara, T., Fujiwara, K., Murakami, Y. & Nakajima, K. 2002. In-situ monitoring system of the position and temperature at the crystal–solution interface. *Journal of Crystal Growth*, 236, 125–131.

[20] Azuma, Y., Usami, N., Ujihara, T., Fujiwara, K., Sazaki, G., Murakami, Y. & Nakajima, K. 2003. Growth of SiGe bulk crystals with uniform composition by utilizing feedback control system of the crystal–melt interface position for precise control of the growth temperature. *Journal of Crystal Growth*, 250, 298–304.

[21] Yildiz, M., Dost, S. & Lent, B. 2005. Growth of bulk SiGe single crystals by liquid phase diffusion. *Journal of Crystal Growth*, 280, 151–160.

[22] Adachi, S., Ogata, Y., Koshikawa, N., Matsumoto, S., Kinoshita, K., Yoshizaki, I., Tsuru, T., Miyata, H., Takayanagi, M. & Yoda, S. 2005. Homogeneous SiGe crystals grown by using the traveling liquidus-zone method. *Journal of Crystal Growth*, 280, 372–377.

[23] Miyata, H., Adachi, S., Ogata, Y., Tsuru, T., Muramatsu, Y., Kinoshita, K., Odawara, O. & Yoda, S. 2007. Crystallographic investigation of homogeneous SiGe single crystals grown by the traveling liquidus-zone method. *Journal of Crystal Growth*, 303, 607–611.

[24] Honda, T., Suezawa, M. & Sumino, K. 1996. Growth and characterization of bulk Si–Ge single crystals. *Japanese Journal of Applied Physics Part 1 – Regular Papers, Short Notes and Review Papers*, 35, 5980–5985.

[25] Dold, P., Barz, A., Recha, S., Pressel, K., Franz, M. & Benz, K. W. 1998. Growth and characterization of $Ge_{1-x}Si_x$ ($x <= 10$ at%) single crystals. *Journal of Crystal Growth*, 192, 125–135.

[26] Labrie, D., George, A. E., Jamieson, M., Obruchkov, S., Healey, J. P., Paton, B. E. & Saghir, M. Z. 2004. Single crystal growth of $Ge_{1-x}Si_x$ alloys using the traveling solvent method. *Journal of Vacuum Science and Technology A*, 22, 962–965.

[27] Azuma, Y., Usami, N., Fujiwara, K., Ujihara, T. & Nakajima, K. 2005. A simple approach to determine preferential growth orientation using multiple seed crystals with random orientations and its utilization for seed optimization to restrain polycrystallization of SiGe bulk crystal. *Journal of Crystal Growth*, 276, 393–400.

[28] Bliss, D., Demczyk, B., Anselmo, A. & Bailey, J. 1997. Silicon–germanium bulk alloy growth by liquid encapsulated zone melting. *Journal of Crystal Growth*, 174, 187–193.

[29] Kadokura, K. & Takano, Y. 1997. Germanium–silicon single crystal growth using an encapsulant in a silica ampoule. *Journal of Crystal Growth*, 171, 56–60.

[30] Usami, N., Nose, Y., Fujiwara, K. & Nakajima, K. 2006. Suppression of structural imperfection in strained Si by utilizing SiGe bulk substrate. *Applied Physics Letters*, 88, 221912.

[31] Usami, N., Nihei, R., Yonenaga, I., Nose, Y. & Nakajima, K. 2007. Application of Czochralski-grown SiGe bulk crystal as a substrate for luminescent strained quantum wells. *Applied Physics Letters*, 90, 181914.

[32] Tsujino, S., Usami, N., Weber, A., Mussler, G., Shushunova, V., Grutzmacher, D., Azuma, Y. & Nakajima, K. 2007. SiGe double barrier resonant tunneling diodes on bulk SiGe substrates with high peak-to-valley current ratio. *Applied Physics Letters*, 91, 032104.
[33] Nihei, R., Usami, N. & Nakajima, K. 2009. Growth of compositionally graded SiGe bulk crystal and its application as substrate with lateral variation in Ge content. *Japanese Journal of Applied Physics*, 48, 115507.

5

Silicon–germanium (SiGe) crystal growth using molecular beam epitaxy

A. SAKAI, Osaka University, Japan

Abstract: This chapter describes molecular beam epitaxy (MBE) techniques as a tool for the formation of nanostructures made with SiGe and related materials. The chapter first describes the growth techniques of MBE, in which two methods using solid and gas sources are introduced. Details of flux control and monitoring are also shown. The chapter then reviews the following topics of nanostructure formation using MBE: Ge islanding phenomena on Si, intrinsically induced nanometer-scale textures in strain-relaxed SiGe, non-equilibrium growth of $Ge_{1-x}Sn_x$ for tensile strained Ge, atomic evolution of $Si_{1-x-y}Ge_xC_y$ thin films on Si, and atomistic reaction between Ge surface and oxygen.

Key words: molecular beam epitaxy (MBE), transmission electron microscopy (TEM), islands, strain relaxation, scanning tunneling microscopy (STM).

5.1 Introduction

Molecular beam epitaxy (MBE) is the most widely used method for depositing semiconductor and metallic materials to form not only single thin films but also multilayers. In this method, because growth occurs under an ultra-high vacuum condition of the order of 10^{-8} Pa, the growth rate can be reduced to a sub-nanometer-scale thickness per second. Thus MBE has great potential for producing several types of nanostructures such as quantum dots, nanowires, multi-stacking atomic layers, etc. In this chapter, we focus on MBE as a tool for the formation of nanostructures made with SiGe and related materials.

The chapter consists of three parts. In the first part, the growth techniques of MBE are explained, in which methods using solid and gas sources are introduced. Details of flux control and monitoring are also shown, which are very important for realizing precise control of nanostructures. In the second part, five topics concerning nanostructure formation using MBE are reviewed. Ge islanding on Si substrates is an attractive system from both scientific and technological viewpoints. Here the interplay between nanoscale surface morphological variation and defect introduction during the Ge islanding is clarified. A strain-relaxed SiGe layer on Si is a model structure for technological applications. Intrinsic nanoscale textures in the MBE-grown SiGe are elucidated and the importance of dislocation engineering is claimed in the second topic. One of the important aspects of MBE is the growth mode which is mostly governed by kinetics rather than thermodynamics due to the low temperature growth. This realizes non-equilibrium growth of $Ge_{1-x}Sn_x$ semiconductor alloys to produce novel strained Ge layers, which will be discussed as a third topic. Due to the reduced pressure environments of MBE, in-situ and real-time diagnostic tools can be readily incorporated into the vacuum system to monitor the dynamics occurring on the growing film surface. As shown in the fourth and fifth topics, in-situ scanning tunneling microscopy (STM) is employed for investigating phenomena at the initial stages of $Si_{1-x-y}Ge_xC_y$ thin film growth on Si(001) surfaces and Ge(001)–oxygen reaction, respectively. For both cases, nanometer- to atomic-scale structural changes on the surfaces can be detected, giving us the fundamental mechanisms of the phenomena. To conclude the chapter, future prospects and trends associated with MBE growth are briefly discussed.

5.2 Techniques

5.2.1 Solid-source method

A silicon MBE system basically consists of four major parts: a vacuum system, a silicon molecular beam source, a dopant source, and a substrate heating system. The vacuum quality relies on the materials used for the chamber, the seals, the pumps, the system design, etc., and must be high enough to grow epitaxial films with high crystalline quality. Because of the high reactivity of silicon atoms, it is essential to reduce the ambient oxygen concentration drastically, so that the base pressure of the MBE chamber is lowered to the range of 10^{-8} Pa. In general, a combination of a cryopump, an ion pump, a titanium sublimation pump and liquid-nitrogen-cooled shrouds is used, as well as an oil-free rough pumping system. A well-controlled baking system at temperatures around 150°C is also a critical requirement to attain ultra-high vacuum (UHV).

As a solid source of Si, an electron gun evaporator is usually used in order to gain a vapor pressure available for practical growth. One representative electron gun evaporator has an electron gun with a 270° electron beam deflector, in which the tungsten filament is hidden from the silicon substrate, minimizing tungsten contamination. The evaporation of Ge is performed using either the electron gun evaporator or a pyrolytic boron nitride (PBN) Knudsen cell (K-cell) which is surrounded by liquid-nitrogen-cooled shrouds to condense unwanted evaporants and improve the vacuum in the sample region. The K-cell is also used for evaporating dopant materials, which are Ga and B for p-type doping, and As and Sb for n-type.

The growth temperature in MBE is generally much lower than that in chemical vapor deposition (details will be described in Chapter 6). However, a process step at fairly high temperature (approximately 1000°C or higher) is required for cleaning a substrate surface prior to depositing materials. There are, in general, two methods of substrate heating for Si MBE: resistive heating and indirect heating. In the former method, electron current passes through a Si substrate so that the Si substrate itself becomes the only heat source. Typically, a rectangular Si substrate is held with tantalum clamps at both ends of the sample with thin tantalum shims. On the other hand, indirect heating is suitable for treating conventional Si wafers with a disk shape. It is mainly performed by using a heater made of a high melting point metal or graphite, which is specially patterned to realize uniform heating. A diffuser made of PBN or SiC-coated carbon enables further uniformity.

5.2.2 Gas-source method

Gas sources of Si and Ge such as SiH_4, Si_2H_6 and GeH_4 are often used for epitaxial growth in the same type of UHV chambers as that used for solid-source MBE. In gas-source MBE, as in solid-source MBE, only the Si molecular beam directly irradiated to the substrate contributes to the growth, since other molecular beams are condensed at the shrouds cooled by liquid nitrogen. The source gas pressure during growth is generally on the order of 10^{-3} Pa. One of the advantages in using a gas source instead of a solid source is the ease of source exchange without breaking the UHV of the MBE system. Another apparent advantage relies on the ability to produce selective epitaxial growth on SiO_2 mask patterned Si substrates (Hirayama *et al.*, 1987), where the growth occurs not on the SiO_2 surface but on the Si surface. This phenomenon raises the potential for creating various types of nanostructures on lithographically patterned Si substrates which are frequently used for the fabrication of electronics and opto-electronics devices.

Growth kinetics of epitaxy on clean Si surfaces have been extensively studied so far using SiH_4 (Suemitsu *et al.*, 1990), Si_2H_6 (Kim *et al.*, 1998), GeH_4 (Murata and Suemitsu, 2004), Ge_2H_6 (Lam *et al.*, 1997), and organosilanes (Foster *et al.*, 1997). In these studies, infrared spectroscopy, temperature-programmed desorption, and electron energy loss spectroscopy are effectively applied to analyze adsorption and desorption kinetics of adspecies on the surfaces. In general, the growth rate of SiGe on Si is dependent on substrate temperature and flow rate of source gases. The growth mainly consists of two modes depending on the temperature. At lower temperatures, the growth mode is in the surface-reaction-limited regime where the growth is limited by the chemical reaction of impinging gases and the growth rate increases with the increase in Ge concentration. On the contrary, at higher temperatures, the growth proceeds in the supply-limited regime where the growth rate saturates and decreases with increasing Ge concentration. These phenomena are observed independently of species of source gases, either monohydride (SiH_4 and GeH_4) or dihydride (Si_2H_6 and Ge_2H_4) sources. Several models are now established that consider kinetic parameters such as the sticking probability of gas species at adsorption sites, hydrogen desorption, and Ge segregation (Robbins *et al.*, 1991; Gu *et al.*, 1996).

5.2.3 Flux control and monitoring

The flux control of Si and Ge deposition is one of the critical issues in obtaining high quality, device-grade thin films. Thus the accuracy of growth rate measurement is a critical factor in the operation of the MBE system. A quartz crystal microbalance is the most commonly used deposition sensor. This utilizes the fact that the quartz oscillation frequency decreases with increase of crystal mass. Since, under constant deposition conditions, the deposited mass on the crystal increases linearly with time, the precise measurement of frequency change over a determined time allows us to obtain the deposition rate. However, the quartz oscillation frequency strongly depends also on temperature. Quartz temperature, therefore, must be stabilized around room temperature to avoid frequency drift.

Precise flux control of deposition has also been achieved by using a mass spectrometer as a rate monitor (Ni *et al.*, 1994). For Si and Ge deposition, mass numbers 28 and 74 are selected as control signals, respectively. The measured partial pressures for individual species need to be converted to deposition rates by calibration measurement such as reflection high-energy electron diffraction (RHEED) intensity oscillations (for details, see below).

There is also a multi-species sensor mainly used in solid-source MBE, which is based on electron impact emission spectroscopy (EIES). EIES is the electron beam excitation of in-flight deposition material to create photons, i.e., luminescence. The created light passes through an optical filter to a photomultiplier tube detector to measure the intensity of emission of the passed wavelength. Then a feedback system generates a signal for controlling the flux of that material. Since the luminescence frequency is determined by atomic energy levels which are unique to each atom species, EIES can be applied to simultaneous co-deposition systems using multiple materials. Deposition sensors based on EIES give not only reliable indications of a growth rate below 0.01 nm/s with an appropriate time response but also thermal immunity suitable for tight feedback control of electron gun deposition sources so that they can be applied to a variety of Si MBE including SiGe and metal silicide materials system growth.

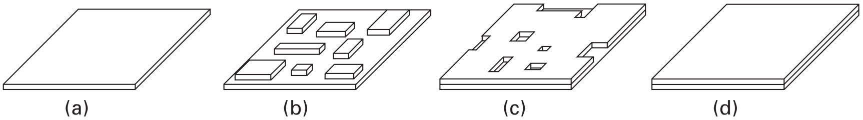
In the case of deposition of low vapor pressure materials which are often evaporated by using electric gun heating, the probability of sticking on the growing film surface tends to unity and thus a measurement of the incoming flux is a true measure of the growth rate. However, this is not necessarily the case for volatile molecules, such as a column V element, which have sticking probabilities depending on their partial pressure, the substrate temperature and the growing surface structure. Furthermore, as mentioned above, calibration measurements are sometimes required for directly converting control signals selected by flux sensors to the growth rate on the film surface. Thus, direct monitoring of the growth rate is always a necessity in MBE and the most probable method for this is the detection of RHEED oscillations during film growth (Harris *et al.*, 1981; Sakamoto *et al.*, 1987).

The electron beam with an accelerating voltage of 10–30 kV at glancing incidence is reflected on the film surface with the diffraction signal coming from only the first few atomic surface layers of the growing film. RHEED intensity oscillation can be explained by equating the changes in intensity of the specular beam with changes in the surface roughness of the growing film. When the surface is atomically smooth without any steps, as shown in Fig. 5.1(a), the reflectivity of the electron beam is a maximum. As a film grows on the smooth surface, two-dimensional (2D) clusters are first formed at random positions on the surface, leading to a decrease in the reflectivity (Fig. 5.1(b)). Approximately half-layer coverage of such 2D clusters, corresponding to the roughest surface, results in the minimum intensity of the beam. The reflectivity again increases as the 2D clusters develop and coalesce with each other, completing one layer growth (Figs 5.1(c) and 5.1(d)). Thus one period of RHEED oscillation directly corresponds to one layer growth of the film, giving a reliable growth rate during MBE. In practice, damped oscillation is often observed, showing that the amplitude gradually decreases with an increase in the growth period. This phenomenon is well interpreted by the occurrence of surface roughening in which the next layers start to develop on the 2D clusters before the completion of the first layer and the surface is statistically distributed over several incomplete atomic levels (Neave *et al.*, 1983).

5.3 Nanostructure formation by molecular bean epitaxy (MBE)

5.3.1 Germanium (Ge) growth on Silicon (Si)

In thin film heteroepitaxy where the film has a lower surface free energy than the substrate but there is large lattice mismatch, films grow in the

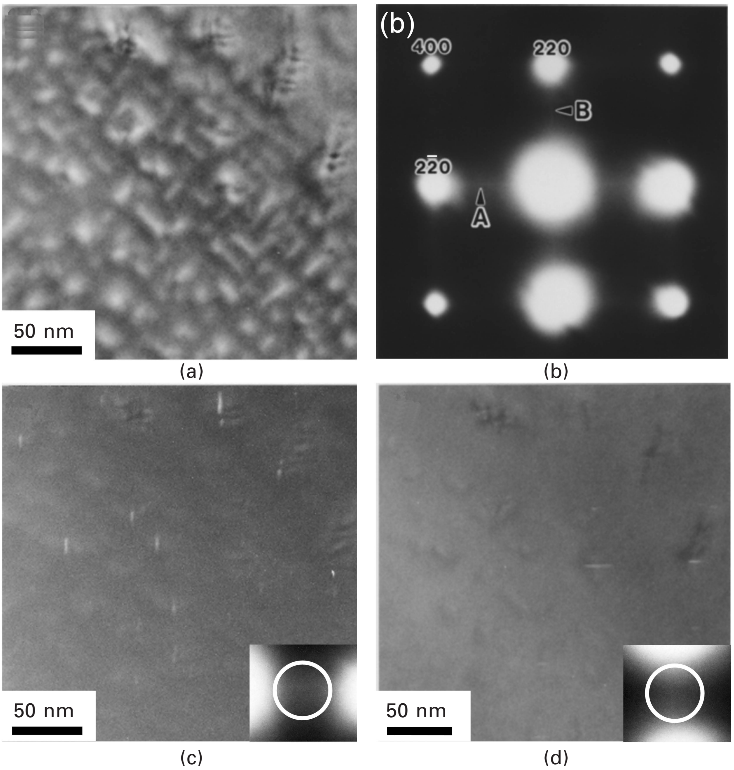


5.1 Schematic illustration of the formation of a single complete layer.

Stranski–Krastanov (SK) mode. Growth of Ge on Si is a typical SK system: about three-monolayer (ML) layer-by-layer growth of Ge is followed by three-dimensional island formation caused by the lattice mismatch where the lattice constant of Ge is approximately 4.2% larger than that of Si. Since defect introduction and film islanding are essential processes for releasing the built-in strain energy, several types of nanostructures appear during film growth. Extensive studies using MBE have been achieved, revealing the early stages of Ge on Si(001) surfaces and showing several types of nanoscale islands that are called hut clusters or faceted islands (Mo *et al.*, 1990; Iwasaki *et al.*, 1991; Sakai and Tatsumi, 1993), dislocation-free or non-faceted islands (Eaglesham and Cerullo, 1990), macro-islands (Sakai and Tatsumi, 1994), pyramidal and dome islands (Medeiros-Ribeiro *et al.*, 1998; Ross *et al.*, 1998). In this subsection, the manner of Ge film evolution on Si(001) surfaces during MBE is explained in detail, especially for the initial stage of nanostructure formation in Ge islanding.

Ge growth on Si proceeds in a complicated manner in which morphological variation of the surface is closely related to strain relief of the evolving film. In order to detect simultaneously both surface morphological evolution of Ge films and defect structures inside the film and/or at the Ge/Si interface, transmission electron microscopy (TEM) can be effectively employed. Figure 5.2(a) is a typical plan-view bright-field TEM image of a sample in which 7-ML Ge was deposited by solid-source MBE at a substrate temperature of 300°C on a Si(001) substrate. The image clearly shows nanoscale rectangular and square structures identified as the hut clusters (faceted islands) that have also been identified by scanning tunneling microscopy (STM) (Mo *et al.*, 1990). Figures 5.2(c) and 5.2(d) are corresponding dark-field images of the same area, taken with the respective extra spots indicated by arrowheads A and B in the transmission electron diffraction (TED) pattern of Fig. 5.2(b), respectively. They exhibit thin defects running along the <110> direction perpendicular to the elongation direction of the spot. A comparison between the bright-field and dark-field images explicitly shows that the defects are situated where the hut clusters coalesce. The atomic structure of the defect can be revealed by cross-sectional TEM observation through the running <110> projection. Figure 5.3 shows a high-resolution (HR) TEM image of the defect together with an atomic model derived from the image simulation (Sakai and Tatsumi, 1993). The defect has a twin structure with a <211> surface normal at the center and the twin region is separated from the matrix by two Σ9 grain boundaries, forming a V-shape.

Such a structural feature of the defect, hereafter called T-Σ9, and the fact that it is formed where two hut clusters coalesce, strongly suggest the nature of a grown-in defect during film evolution, rather than an extrinsically introduced defect such as misfit dislocations. Following this idea, the formation mechanism of T-Σ9 is deduced. Here we consider that a pair of

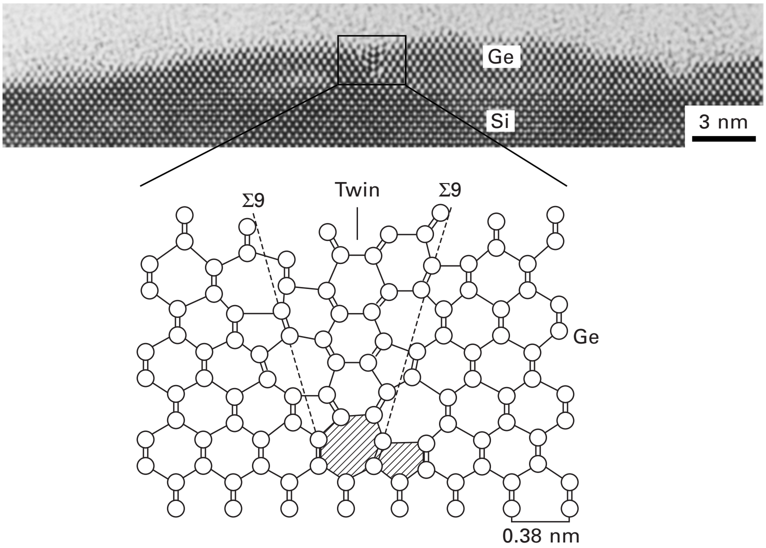


(a) (b)

(c) (d)

5.2 (a) Plan-view bright-field TEM image of a 7-ML Ge sample grown at 300°C showing hut clusters. (b) Corresponding TED pattern showing two types of extra spots indicated by the arrowheads A and B. (c) and (d): Plan-view dark-field TEM images of the same area taken with the 'A' and 'B' extra spots circled in the inset, respectively.

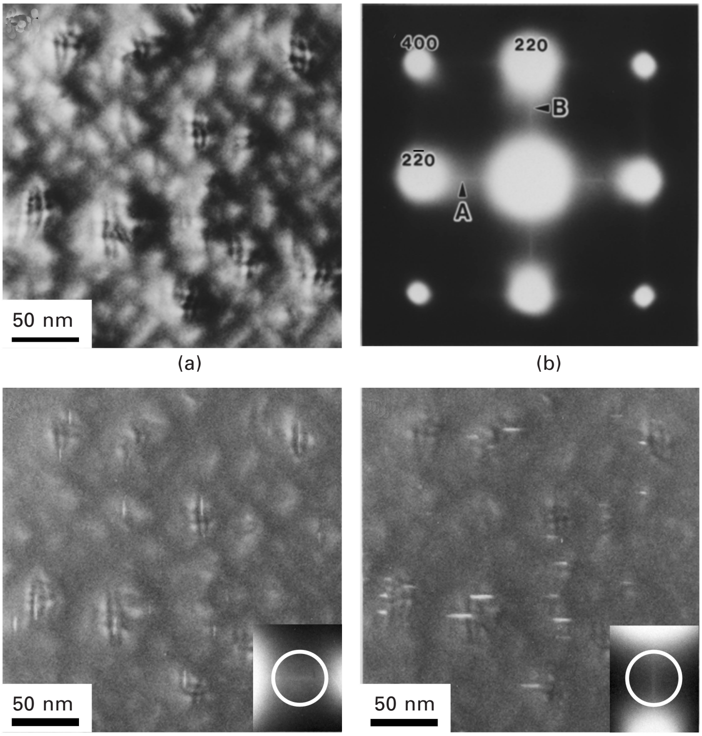
five-membered and seven-membered atomic rings (P5-7AR) at the cusp of T-Σ9, seen as the shaded zone in the atomic model in Fig. 5.3, must nucleate at the first stage of defect formation. The most probable candidate causing this nucleation is the surface dimer structure, which is essentially identical to that of the five-membered atomic ring. Compressive stress caused by the lattice mismatch between Ge and Si likely stabilizes the surface dimer bonds on the surface of the Ge growing film (Men *et al.*, 1988). Furthermore, it is expected that more compressive stress is accumulated at the sites where the hut clusters coalesce, since a combination of the slope of the hut clusters directly means a loss of the step contribution to stress compensation. In view of this, the dimer bond at the coalescence site becomes less reactive, preserving the five-membered atomic ring structure during growth; this gives rise to a reduction of the nucleation barrier for P5-7AR formation. Once



5.3 Cross-sectional HRTEM image of the 7-ML Ge sample grown at 300°C showing the T-Σ9 defect between two hut clusters in the <110> projection. An atomic model of the defect is also shown, in which the shaded region at the cusp of the defect shows a pair of five- and seven-membered atomic rings.

P5-7AR nucleates on the growth surface, diffusing atoms are atomically incorporated, easily forming twin and Σ9 boundary structures on it, since the Σ9 boundary also consists of chained P5-7ARs.

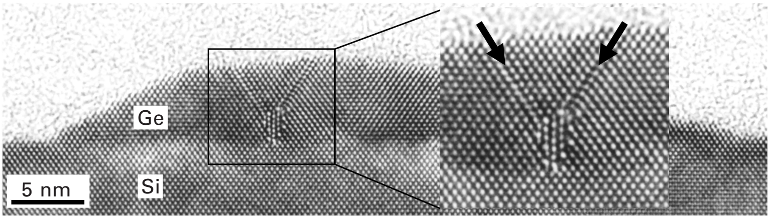
Further deposition of Ge leads to the onset of macro-island formation on the hut clusters. In Fig. 5.4(a), a bright-field TEM image of a 12-ML Ge sample shows macro-islands exhibiting moiré fringes as well as hut clusters. Dark-field images of the corresponding region are shown in Figs 5.4(c) and 5.4(d), which were taken in the same manner as in Fig. 5.2. A comparison between the bright-field and the dark-field images reveals a remarkable one-to-one correspondence between the formation sites of macro-islands and defects. This indicates that the macro-islands prefer to grow in the vicinity of T-Σ9, meaning that nucleation for macro-island formation is predominantly heterogeneous. On the contrary, there is another report on such defect-mediated macro-island formation observed by in-situ UHVTEM (Hammer *et al.*, 1996). Although the phenomenon has been confirmed for Ge islanding using a GeH_4 gas source on thin foil Si(001) substrates, it has been shown that the macro-island formation is mediated not by T-Σ9 but by 90° dislocations formed due to the coalescence of hut clusters. This difference can be understood by considering the displacement vector of T-Σ9. As clearly seen in Fig. 5.5, a T-Σ9 buried under the macro-island is



(a) (b)

(c) (d)

5.4 (a) Plan-view bright-field TEM image of a 12-ML Ge sample grown at 300°C showing macro-islands as well as hut clusters. (b) Corresponding TED pattern showing two types of extra spots indicated by the arrowheads A and B. (c) and (d): Plan-view dark-field TEM images of the same area taken with the 'A' and 'B' extra spots circled in the inset, respectively. All defect images overlap with the moiré fringes of the macro-islands.



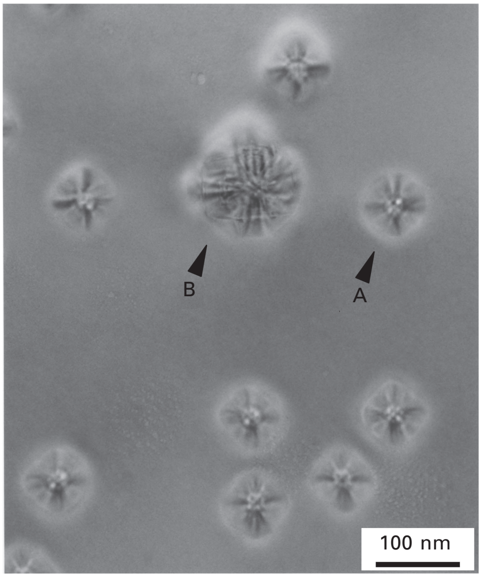
5.5 Cross-sectional HRTEM image showing a macro-island in the <110> projection. A T-∑9 accompanied by two intrinsic stacking faults indicated by arrows is observed.

frequently accompanied with two intrinsic stacking faults (see arrows) and the total displacement vector $\mathbf{b}_{\text{T-}\Sigma9}$ is

$$\mathbf{b}_{\text{T-}\Sigma9} = \frac{a}{6}[21\bar{1}] + \frac{a}{6}[\bar{1}\bar{2}1] + \frac{a}{2}[1\bar{1}0] \qquad 5.1$$

Therefore, nucleation of two Shockley-partial dislocations during growth results in the annihilation of the faults and the formation of a 90° dislocation with a Burgers vector of $a/2\langle 110\rangle$ type. This consideration leads to a conclusion that the nucleation process of the 90° dislocation is substantially identical to that of T-Σ9 and the polytype formation of the defects is determined by the Ge atom stacking sequence which might depend on the growth conditions.

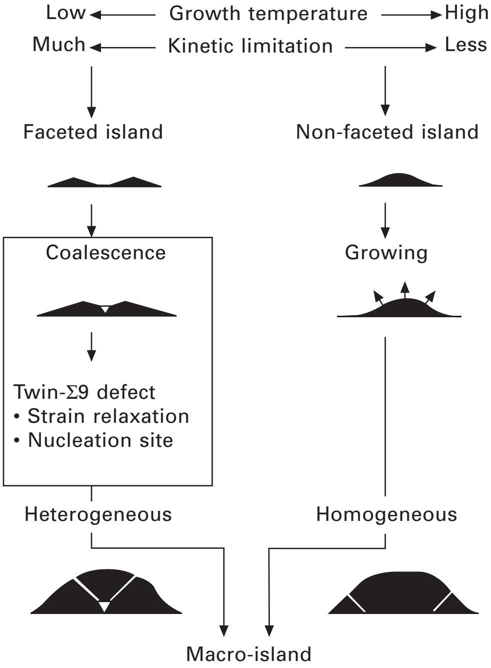
The above defect-mediated macro-island formation is dominant for growth temperatures up to ~450°C. However, at temperatures higher than ~550°C, the Ge islanding tends to proceed in a homogeneous mode. Figure 5.6 is a typical plan-view bright-field image of a 7-ML Ge sample grown at 600°C. Macro-islands with lower number density than that of Fig. 5.4(a) are formed but neither hut clusters nor T-Σ9 are present at any stages of growth. The observed smaller macro-islands (one of which is denoted by A) are dislocation-free and correspond to the 'pyramids'. On the other hand, the larger island denoted by B, referred to as a 'dome', contains defects. Thus macro-island formation at this temperature would proceed as follows. By depositing an



5.6 Plan-view bright-field TEM image of a 7-ML Ge sample grown at 600°C. Smaller dislocation-free islands denoted by A and a larger island containing defects denoted by B are observed.

increasing amount of Ge, non-faceted, initially dislocation-free islands are continuously grown without forming T-Σ9. After a certain amount of Ge is deposited, defects are newly introduced into the grown island, resulting in complete strain relaxation.

Dominance of either a homogeneous or a heterogeneous mode in the island nucleation stage should be explained by considering the thermodynamic driving force for islanding and its kinetic limitations. The mechanism of Ge film evolution on Si(001) can be explained according to Fig. 5.7. At high temperatures, with less kinetic limitation, a simple growth process occurs in which an initially small dislocation-free, non-faceted island develops into a larger macro-island (see the right-hand route in the figure). Therefore, the nucleation is homogeneous. On the other hand, lowering the growth temperature kinetically limits the driving force for macro-island formation. In this situation, the system might be driven into T-Σ9 formation due to the coalescence of hut clusters, instead of direct formation of the macro-island on the surface (see the left-hand route in the figure). T-Σ9 primarily plays two important roles in Ge film evolution. First, it partially releases strain energy of the evolving film before macro-island formation. Since a T-Σ9 has close-packed {111} planes parallel to the {110} plane of the substrate and


Low
Growth temperature
High
Much
Kinetic limitation
Less
Faceted island
Non-faceted island
Coalescence
Growing
Twin-Σ9 defect
• Strain relaxation
• Nucleation site
Heterogeneous
Homogeneous
Macro-island

5.7 Schematic diagram explaining a mechanism of macro-island formation in the Ge/Si(001) system.

a V-shaped morphology, the compressive stress due to lattice mismatch is gradually relaxed as the film grows (LeGoues *et al.*, 1989). Second, it directly mediates macro-island formation by acting as a preferential nucleation site. It is well known that the energy barrier for nucleation is reduced by the free energy of a defect (Christian, 1982, p. 448). Therefore macro-island formation easily proceeds in a heterogeneous nucleation mode, even with the kinetically limited driving force, so as to decrease the extra free energy created by the defect.

The interplay between nanoscale surface morphological variation and defect introduction in Ge films has been clarified. Macro-island formation proceeds in both heterogeneous and homogeneous mode, depending on the growth temperature. In the heterogeneous mode, particular defects caused by coalescence of hut clusters predominantly mediate macro-island formation. A mechanism of macro-island formation of Ge on Si(001) surfaces can be explained by a combination of the thermodynamics driving force for islanding and its kinetic limitations.

5.3.2 Nanoscale textures in SiGe/Si systems

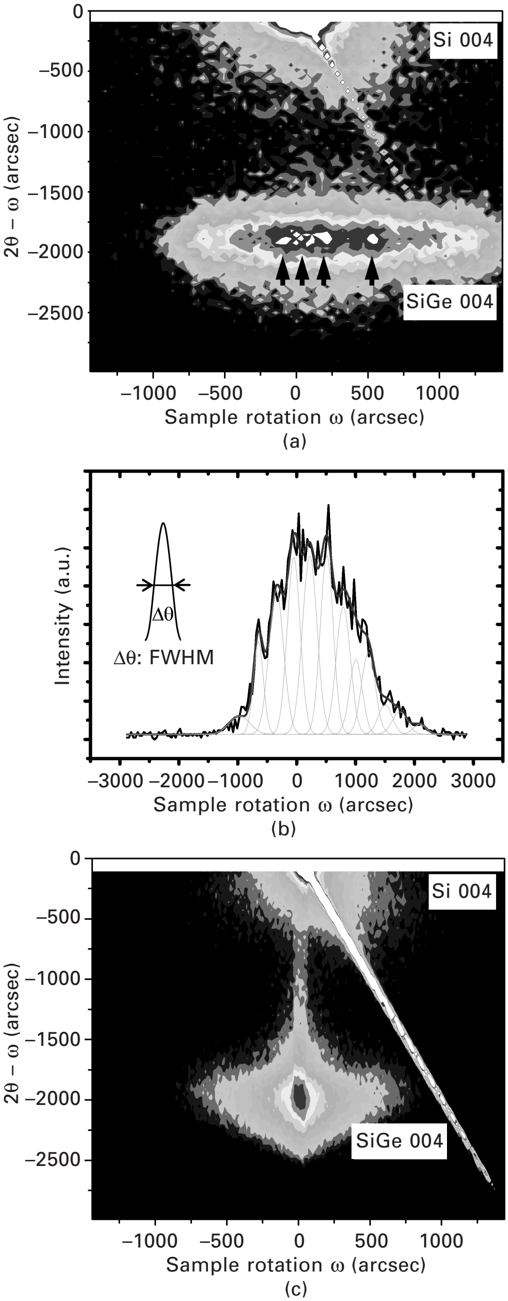
A high-quality strain-relaxed SiGe layer grown on Si as a virtual substrate offers one of the crucial applications of Si-based high-speed metal-oxide-semiconductor field effect transistors (MOSFETs) with tensile strained Si channels. As explained in more detail in Chapter 7, formation of the virtual substrates is based on the strain relaxation process of SiGe layers. A MBE grown SiGe film on a Si(001) substrate is a prototype showing how compressive strain in the film relaxes during growth and how defects, especially misfit dislocations, are introduced into the film and the SiGe/Si interface. Owing to the diamond lattice structure of SiGe and Si and the <110>/{111} slip system for plastic deformation, glide dislocations, i.e., 60° dislocations, are predominantly introduced into the SiGe/Si(001) interface for strain relaxation. Although the Burgers vector of the 60° dislocation has an edge component parallel to the interface, which mainly contributes to relaxing misfit strain, it has also a screw component and components non-parallel to the interface, which simultaneously induce non-isotropic strain relaxation. Thus the crystallographic tilting and rotation occur in the relaxed SiGe layer, which results in the deteriorating performance of devices. In this subsection, we focus on the nanometer-scale crystalline textures inevitably generated in MBE-grown strain-relaxed SiGe layers by the introduction of the 60° dislocation. Dislocation engineering for realizing high crystalline quality SiGe layers is also discussed.

X-ray microdiffraction (XRMD) is one of the powerful tools by which local strain in a sub-micrometer-scaled area of processed materials can be non-destructively characterized. The hard X-ray undulator beam line of the

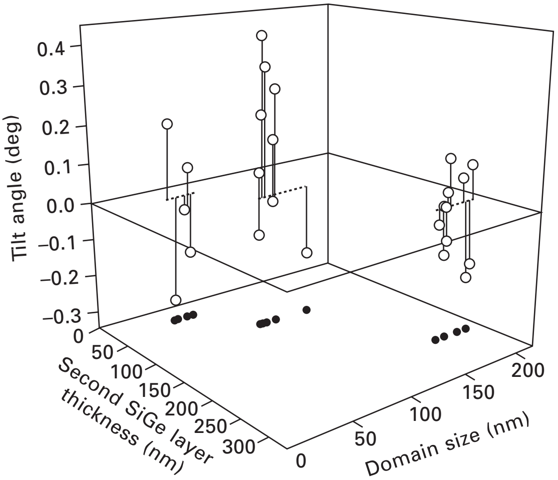
Super Photon ring 8 GeV (SPring-8) in Japan possesses a focused beam area of $0.5 \times 1.6\ \mu m^2$ using a zone plate with narrow slits (Takeda *et al*., 2006). Using this technique, a variation of local lattice spacing and tilting has been analyzed for two types of MBE-grown SiGe layers on Si(001) substrates (Mochizuki *et al*., 2006). One was a SiGe layer grown by the two-step strain-relaxation procedure (Sakai *et al*., 2001a), referred to as sample A. In this procedure, after growing the first $Si_{0.7}Ge_{0.3}$ layer with a Si cap layer on a Si(001) substrate at 400°C, the sample was annealed at 600°C to form a network of 60° dislocations running along two orthogonal <110> directions at the SiGe/Si interface. The second $Si_{0.7}Ge_{0.3}$ layer was then grown up to a thickness ranging from 30 to 250 nm at 600°C to promote strain relaxation. Another SiGe layer was prepared by solid-phase intermixing of Si with a preformed Ge film on Si(001) with high-temperature annealing, referred to as sample B. In this case, a 35-nm-thick Ge layer was epitaxially grown on a Si(001) substrate at 200°C and a 17-nm-thick amorphous Si layer was subsequently deposited at room temperature. The sample was then annealed at 1100°C in a N_2 ambient.

Figure 5.8(a) shows an XRMD two-dimensional reciprocal space map (2DRSM) around the Si 004 and SiGe 004 diffraction spots of sample A with a 100-nm-thick second SiGe layer. It is found that the SiGe(004) diffraction peak broadens along the ω (rotation angle) direction. Furthermore, as indicated typically by arrows in the figure, several discrete peaks can be observed in the broadened SiGe 004 peak. Such discrete peaks are never seen in conventional macroscopic X-ray diffraction (XRD) maps taken with a millimeter-sized beam.

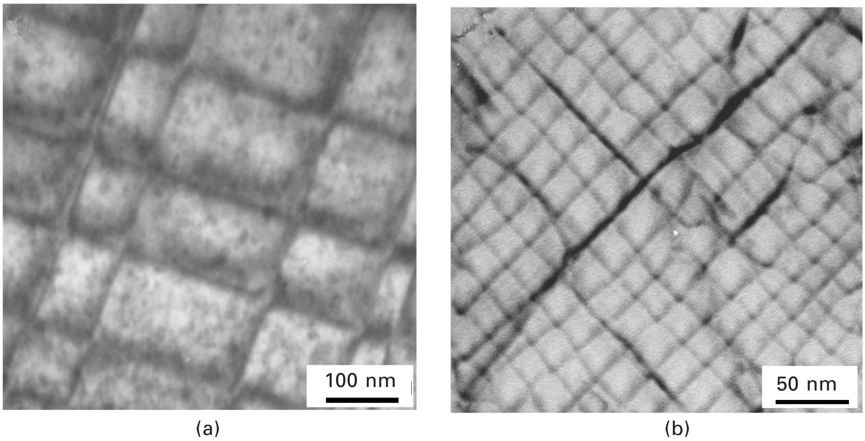
Detailed crystalline textures of sample A have been analyzed from the data. Figure 5.8(b) is a cross-sectional profile at $2\theta = -1890$ arcsec of the SiGe 004 diffraction peak of Fig. 5.8(a), showing a complex fine structure which consists of multi-peaks. This indicates that there are several crystal domains tilting individually at different angles in the observed region, of approximately $1.0 \times 9.1\ \mu m^2$, and each discrete diffraction peak comes from a nanometer-sized crystal domain. Tilt angles and in-plane sizes of the domains estimated from individual discrete peaks are shown in Fig. 5.9 as a function of the second SiGe layer thickness. The absolute values of measured tilt angles for the domains range from 0.00° to 0.42° with respect to the [001] axis of the Si substrate. It can be seen that the tilt angle once increases and then decreases while the average domain size monotonically increases with increase in the second layer thickness. As shown in a plan-view TEM image of Fig. 5.10(a), in the two-step strain relaxation procedure, 60° misfit dislocations are introduced at the SiGe/Si(001) interface and the density increases with increase in the second layer thickness. Because of the crystallographic tilting caused by individual 60° dislocations, the increase in the density enhances the tilting. However, at the same time, the opportunity



5.8 (a) XRMD-2DRSM around the Si 004 and SiGe 004 diffraction spots of sample A with the second SiGe layer thickness of 100 nm. Arrows indicate discrete peaks only observed by the microdiffraction. (b) Cross-sectional profile at $2\theta = -1890$ arcsec of the SiGe 004 peak of sample A shown in (a). (c) XRMD-2DRSM of sample B showing a single sharp peak of SiGe 004 diffraction.



5.9 Measured tilt angles of individual domains and crystal domain sizes of the SiGe layer in sample A as a function of the second SiGe layer thickness.



(a)

(b)

5.10 Plan-view TEM images of (a) 60° dislocations formed at the SiGe/Si(001) interface in sample A, and (b) pure-edge dislocation network at the SiGe/Si(001) interface in sample B. In (b), more densely and regularly aligned dislocations than those in (a) are observed.

of introducing the other type of 60° dislocations which can counteract the pre-existing tilting also increases as the layer thickness increases. This mechanism might account for the observed decrease of the tilt angle and the increase in the domain size.

On the contrary, as shown in Fig. 5.8(c), a single sharp peak is seen in the XRMD-2DRSM of sample B, indicating that the SiGe layer hardly has a mosaic structure. Figure 5.10(b) shows the SiGe/Si(001) interface of sample B where a network of regularly and densely aligned pure-edge dislocations, not the aforementioned 60° dislocation network, contributes to strain relaxation of the SiGe layer. Homogeneous strain distribution with reduced mosaicity can be achieved due to the pure-edge dislocation with the Burgers vector having a single edge component along the in-plane <110> direction (Sakai *et al.*, 2005).

In MBE methods, many sophisticated techniques to grow strain-relaxed SiGe layers have been proposed so far but, in any case, dislocation engineering is of great significance for achieving high crystalline quality. The motivation for this engineering comes from the issues caused by the character of misfit dislocations which predominantly contribute to strain relaxation. The method based on strain relaxation with the pure-edge dislocation network opens up a novel approach to realizing the growth of SiGe layers with truly high crystalline quality.

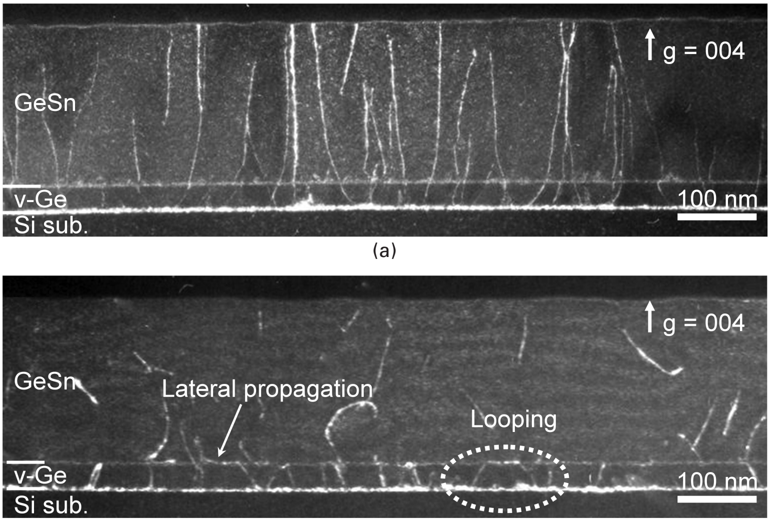
5.3.3 GeSn growth on Si for strained Ge

Introduction of Ge, instead of Si, to the channel of MOSFET is one of the promising solutions for several problems associated with the recent tremendous down-scaling of MOSFET. Ge has approximately twice the mobility for electrons and four times the mobility for holes than does Si. It is therefore expected that both higher drive current and lower power consumption are simultaneously achieved in the operation compared to Si-MOSFET. In addition, a theoretical prediction has made strained Ge more attractive than bulk-Ge; both hole and electron mobilities are significantly increased when the Ge(001) layer has in-plane biaxially tensile strain (Fischetti and Laux, 1996). Although the difference in thermal expansion coefficient between Ge and Si has been utilized to induce tensile strain into Ge films (Ishikawa *et al.*, 2003), heteroepitaxy with precise lattice parameter engineering is greatly preferable from a viewpoint of device application. In heteroepitaxy, a buffer layer plays a crucial role in accommodating mismatch strain with respect to Si substrates and inducing global strain into the Ge channel. $Ge_{1-x}Sn_x$ alloy having a larger lattice constant than that of bulk Ge is one of the most attractive candidates for the buffer layer (Fang *et al.*, 2007; Takeuchi *et al.*, 2007, 2008a, 2008b). This subsection focuses on MBE growth of $Ge_{1-x}Sn_x$ buffer layers on Si substrates for tensile-strained Ge layers. Such group IV material-based heteroepitaxy is evaluated on the basis of precise structural analysis using XRD-2DRSM and TEM.

Because it is clear that the large lattice mismatch between $Ge_{1-x}Sn_x$ and Si can be relieved by the insertion of a Ge layer between them, the use of a virtual Ge substrate (v-Ge) is an effective way to obtain strain-relaxed

$Ge_{1-x}Sn_x$ on Si substrates. For growing a v-Ge, which is a fully strain-relaxed Ge film on Si(001), a 40-nm-thick Ge layer is deposited at 200°C on a clean surface of a Si(001) substrate, followed by ex-situ rapid thermal annealing at 700°C for 1 min in a N_2 ambient to relax the strain of Ge. Most of the strain in the v-Ge is relaxed by the pure-edge dislocation network formed at the Ge/Si interface. $Ge_{1-x}Sn_x$ layers are then grown by MBE at 200°C on the v-Ge after cleaning the surface of v-Ge. The dark-field cross-sectional TEM image of Fig. 5.11(a) shows that an epitaxial $Ge_{0.974}Sn_{0.026}$ layer is grown with threading dislocations which come mostly from those pre-existing in the v-Ge through the $Ge_{0.974}Sn_{0.026}$/v-Ge interface. Such dislocation morphology changes drastically after ex-situ post-deposition annealing (PDA) at 600°C for 10 min in a N_2 ambient, as shown in Fig. 5.11(b). Looping and lateral propagation of the dislocations frequently occur at the $Ge_{1-x}Sn_x$/v-Ge interface and curved dislocations are formed in the layer. Thus the PDA effectively promotes the movement of threading dislocations as well as reduction of threading dislocation density.

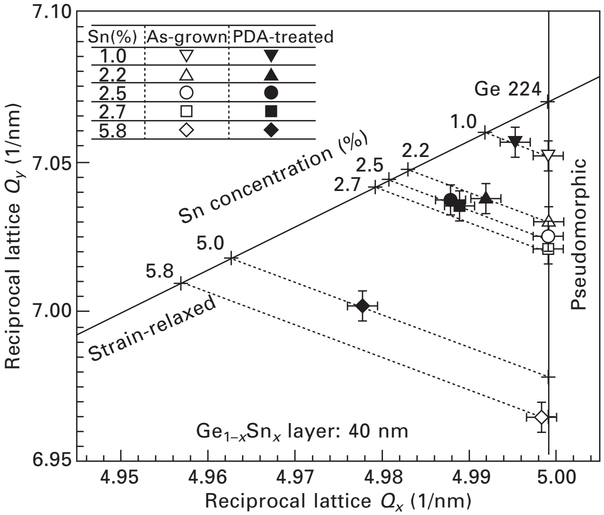
Figure 5.12 depicts $Ge_{1-x}Sn_x$ 224 reciprocal lattice points measured from XRD-2DRSM for $Ge_{1-x}Sn_x$ layers with various Sn concentrations on v-Ge, showing strain relaxation and Sn precipitation behaviors. In the figure, open (filled) symbols represent as-grown (PDA-treated) samples and dotted lines represent the trajectory of the diffraction peak when the strain relaxation



(a)

(b)

5.11 Cross-sectional dark-filed TEM images of $Ge_{0.974}Sn_{0.026}$ layers on v-Ge with a diffraction vector **g** = 004; (a) as-grown and (b) PDA-treated samples. Lateral propagation at the GeSn/v-Ge interface and looping of dislocations are observed.



5.12 $Ge_{1-x}Sn_x$ 224 reciprocal lattice points of single $Ge_{1-x}Sn_x$ layers with constant Sn content on v-Ge measured by XRD-2DRSM. The Sn contents (film thicknesses) of $Ge_{1-x}Sn_x$ were 1.0% (40 nm), 2.2% (210 nm), 2.5% (360 nm), 2.7% (40 nm) and 5.8% (40 nm). The solid lines represent the calculated reciprocal lattice sites for pseudomorphic $Ge_{1-x}Sn_x$ crystals and strained-relaxed cubic $Ge_{1-x}Sn_x$ crystals. The dotted lines represent the calculated reciprocal lattice sites of strain-relaxed $Ge_{1-x}Sn_x$ crystals complying with Vegard's law. Open symbols indicate $Ge_{1-x}Sn_x$ 224 reciprocal lattice points for as-grown samples and solid symbols for PDA-treated samples.

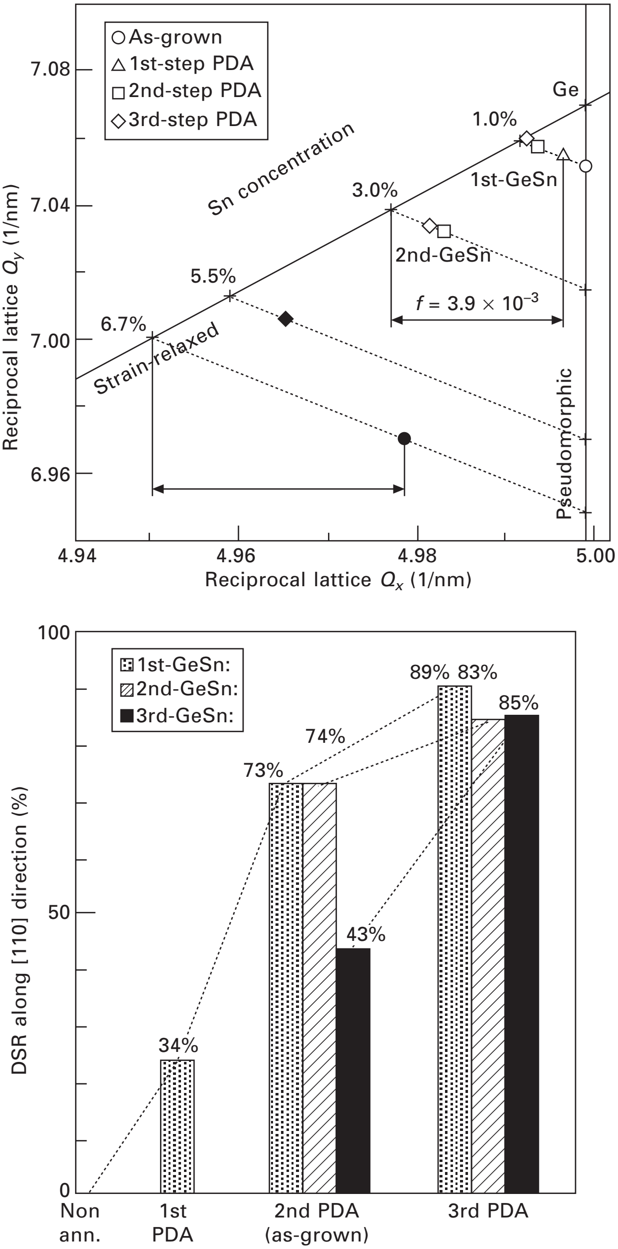
occurs without any variation of substitutional Sn concentration. In the case of Sn concentration of 2.5% and less, two points of as-grown and PDA samples lie on the same line, meaning no concentration variation after PDA. On the other hand, when Sn concentration exceeds 2.5%, a remarkable reduction of substitutional Sn concentration, meaning Sn precipitation, is observable. This analysis shows that there is a critical misfit strain, f_c, for Sn precipitation at a given thickness, i.e., $f_c = 3.7 \times 10^{-3}$ in this case.

As mentioned above, in the case of direct growth of $Ge_{1-x}Sn_x$ layers with high Sn concentration, Sn precipitation markedly occurs. In order to suppress such precipitation and increase the Sn concentration, growth of compositionally step-graded (CSG) $Ge_{1-x}Sn_x$ buffer layers on v-Ge is one of the most effective techniques. By taking f_c into account, one example of the CSG $Ge_{1-x}Sn_x$ buffer layer is designed, which has a structure of $Ge_{0.933}Sn_{0.067}/Ge_{0.97}Sn_{0.03}/Ge_{0.99}Sn_{0.01}$/v-Ge with misfit strains for the bottom two interfaces less than f_c and that for the top one more than f_c. Following this stacking structure, a CSG $Ge_{1-x}Sn_x$ buffer layer was grown by MBE in which PDA was performed after each layer growth, and finally a strained Ge epilayer was grown on top of the CSG layer. $Ge_{1-x}Sn_x$ 224 reciprocal lattice points measured from XRD-2DRSM for this sample are shown in Fig. 5.13, indicating the strain relaxation behavior of each $Ge_{1-x}Sn_x$ layer. Histograms are also shown for degree of strain relaxation (DSR) along the in-plane [110] direction with respect to the pseudomorphic value in each $Ge_{1-x}Sn_x$ layer at respective PDA steps. As the PDA step proceeds, strain relaxation also proceeds without any variation of substitutional Sn concentration for the bottom two layers. For the second layer, the misfit value was measured to be 3.9×10^{-3}, which is close to the value of f_c. On the other hand, reduction of the Sn concentration is apparent in the third $Ge_{1-x}Sn_x$ layer which has the misfit strain measured to be 5.7×10^{-3}, much higher than f_c. Thus the Sn precipitation is approximately determined by the critical misfit strain and a much higher degree of strain relaxation–more than 80%–is achievable in the CSG growth method.

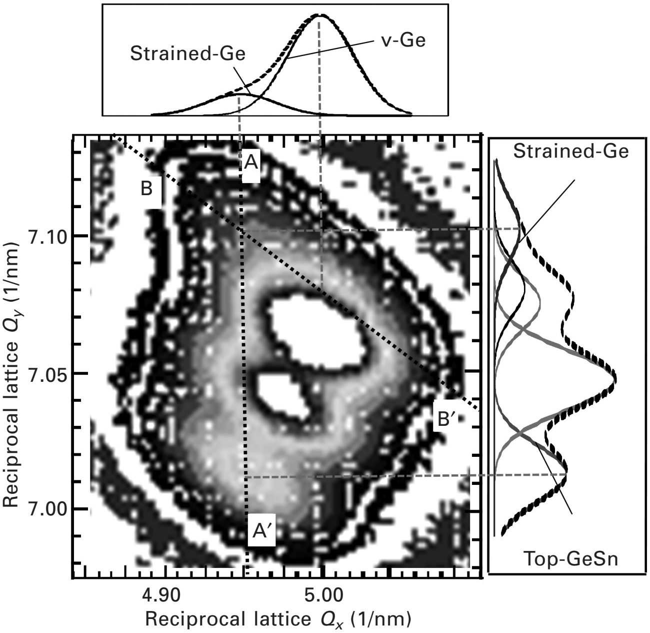
An XRD-2DRSM for 224 reflections taken from the same sample is shown in Fig. 5.14. Diffraction intensity profiles along the lines AA′ and BB′ and deconvolution of the profile indicating diffraction peaks associated with the respective layers are also shown in the figure. The lattice constant of the strained Ge epilayer can be derived as follows. From the profile of AA′ along the [004] direction through the top $Ge_{1-z}Sn_z$ peak, a reciprocal lattice of $Q_y = 0.7103\ \text{nm}^{-1}$ for the strained Ge, which corresponds to the out-of-plane lattice constant of $a_{\perp \text{strained Ge}} = 0.5631$ nm, is obtained. In addition, from the profile of BB′ through the v-Ge peak to the Q_y of the strained Ge, we obtain a reciprocal lattice of $Q_x = 0.4965\ \text{nm}^{-1}$ for the strained Ge, which corresponds to the in-plane lattice constant of $a_{|| \text{strained Ge}} = 0.5696$ nm. Considering the standard error of these reciprocal lattice points for the strained Ge peak in the profile BB′, the in-plane strain value can be estimated to be 0.68 ± 0.03%.

Figure 5.15 shows a corresponding cross-sectional TEM image of the same sample. Observed misfit dislocations originating from pre-existing threading dislocations and lying at each interface surely contribute to strain relaxation of each $Ge_{1-x}Sn_x$ layer. In the strain relaxation process, threading dislocations are often pinned by some reactions between dislocations during PDA. Such pinned dislocations can neither propagate at the interface nor contribute strain relaxation of the layers. However, in the CSG case, the upper layer takes over the pinned threading dislocations in the lower one and the dislocations can move again at the interface during PDA due to misfit stress between the upper and lower layers. This process leads to higher DSR of the stacked $Ge_{1-x}Sn_x$ layers, realizing higher tensile strain in the Ge epilayer.

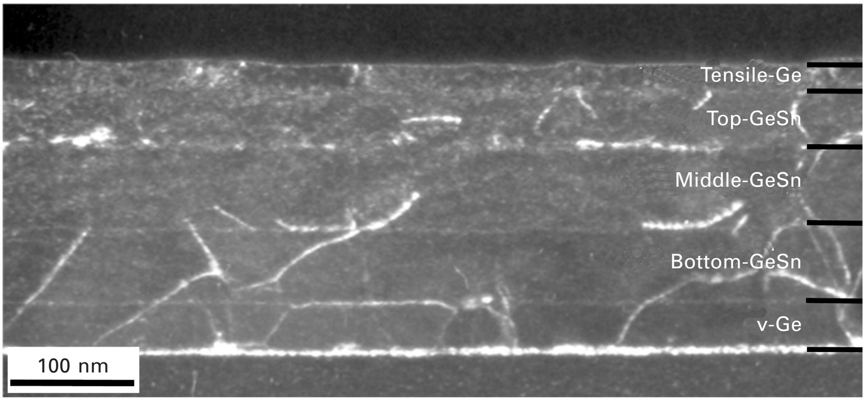
Employment of v-Ge and the CSG growth method realizes high DSR and high substitutional Sn concentration in the $Ge_{1-x}Sn_x$ buffers, both of which allow us to obtain highly tensile-strained Ge epilayers. Dislocation



5.13 $Ge_{1-x}Sn_x$ 224 reciprocal lattice points measured by XRD-2DRSM showing the Sn content variation and the strain relaxation behavior of each $Ge_{1-x}Sn_x$ layer in the Ge/$Ge_{0.933}Sn_{0.067}$/$Ge_{0.970}Sn_{0.03}$/$Ge_{0.99}Sn_{0.01}$/v-Ge structure. Histograms are also shown for DSR along the in-plane [110] direction with respect to the pseudomorphic value in each $Ge_{1-x}Sn_x$ layer at the respective PDA steps.



5.14 XRD-2DRSM for 224 reflections of the same sample as that shown in Fig. 5.13. A diffraction intensity profile along the line AA′ is shown at the right of the figure, and a profile along the line BB′ at the top of the figure. The observed profiles (broken curves) can be deconvoluted into five (two) peaks in the AA′ (BB′) profile. The peaks from strained Ge, top-Ge$_{1-z}$Sn$_z$ and v-Ge are labelled Strained-Ge, Top-GeSn and v-Ge, respectively.



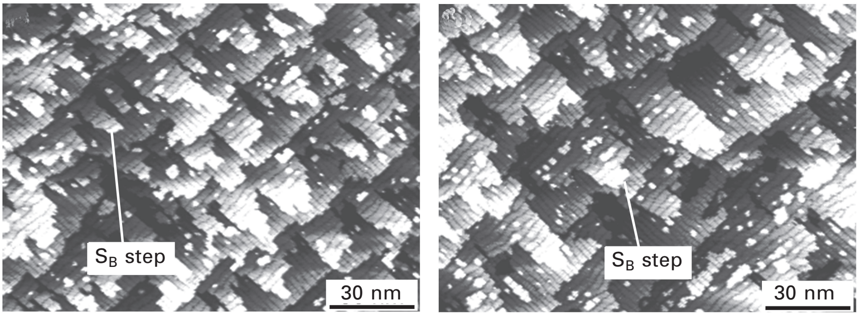
5.15 Cross-sectional dark-field TEM image showing dislocation morphology and interfaces of the sample shown in Fig. 5.14. The image was taken by the weak-beam method at **g**/3**g** with **g** = 004.

propagation at the heterointerface and strain control in the buffer layer is key to tailoring novel channel materials for mobility enhancement solutions.

5.3.4 SiGeC growth on Si

Introduction of C into Si and SiGe films has attracted increasing interest for the application of these films to electronic and optoelectronic devices made with group IV semiconductor materials. Carbon atoms substitutionally introduced into the SiGe lattice sites give rise to substantial decrease of the lattice parameter and reduction of strain in the SiGe layers on Si substrates because of the different covalent radii of Si, Ge and C of 0.117 nm, 0.122 nm and 0.077 nm, respectively. According to a simple calculation using Vegard's law, the compressive strain of 8.2% Ge is compensated by the tensile strain of 1% C in a pseudomorphic $Si_{1-x-y}Ge_xC_y$ film on Si(001) using a linear interpolation between Si, Ge and C (Iyer *et al*., 1992) and the compensation for 9.4% Ge is achieved by 1% C when using the lattice constant of SiC instead of C (Osten *et al*., 1994), although such estimation of the lattice constant requires careful consideration on the basis of theoretical models (Kelires, 1997; Berti *et al*., 1998). While a wide variety of electronic and structural properties brought by $Si_{1-x-y}Ge_xC_y$ films provide a new degree of freedom to electronics technologies (Rücker *et al*., 1999; Oda *et al*., 2003), substitutional incorporation of C exceeding a few percent is very difficult to achieve. One dominant reason is the significantly small thermal equilibrium solubilities of C into Si and Ge, of the order of 10^{17} and 10^8 atoms/cm^3, respectively (Madelung, 1996; Okinaka *et al*., 2003). In general, the crystalline quality of the films is degraded with increasing C fraction and exhibits non-planar morphology, SiC polytype precipitates, and extended defects such as stacking faults and dislocations. Therefore, film growth techniques such as MBE are now widely employed, in which the growth mode is governed not by thermodynamics but by kinetics. In this subsection, we focus on the phenomena on an atomic scale during the initial growth process of $Si_{1-x-y}Ge_xC_y$ thin films on Si(001) surfaces. In-situ STM clarifies mechanisms of the film evolution and the behavior of C atoms during MBE.

First, we look at the difference of the initial surface morphology between MBE-grown SiGe and SiGeC films on Si(001). Figures 5.16(a) and 5.16(b) compare STM images of $Si_{0.5}Ge_{0.5}$ and $Si_{0.489}Ge_{0.489}C_{0.022}$ films at a thickness of 8 ML. Both surfaces exhibit a $(2 \times n)$ reconstructed surface and the undulation of the S_B step (Chadi, 1987) forming elongated domains. An apparent difference between the $Si_{1-x}Ge_x$ and $Si_{1-x-y}Ge_xC_y$ film surfaces is reflected in the separation between adjacent domains, which is measured perpendicular to the elongated direction. The domains on the $Si_{1-x-y}Ge_xC_y$ film surface have a longer separation on average than those on the $Si_{1-x}Ge_x$



(a) (b)

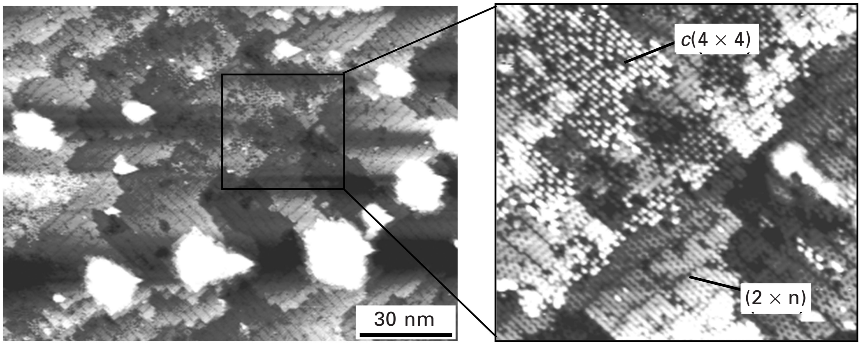
5.16 STM images of the surface structures of (a) $Si_{0.5}Ge_{0.5}$ and (b) $Si_{0.489}Ge_{0.489}C_{0.022}$ films at 8 ML thickness. Both films were grown on Si(001) substrates at 600°C. In both cases, the (2 × *n*) reconstruction is observed as stripes on the terraces and S_B step undulation forms elongated domains.

film surface. This separation, l, can be correlated with the elastic relaxation energy, E, in the form of

$$E = -\frac{C}{l}\ln\left(\frac{l}{\pi a}\right) \tag{5.2}$$

where C is a constant and a is the surface lattice constant, meaning the contribution of surface steps to the strain energy relaxation in the film (Alerhand *et al.*, 1998). Comparison of the E values obtained from an average separation of ~18 nm for $Si_{1-x-y}Ge_xC_y$ to ~12 nm for the $Si_{1-x}Ge_x$ film leads to approximately a 22% reduction in the step contribution to the strain energy relaxation in the $Si_{1-x-y}Ge_xC_y$ film, relative to that in the $Si_{1-x}Ge_x$ film. In other words, C atoms incorporated into the film explicitly play a role in compensating the compressive strain.

In the case of high C content in the $Si_{1-x-y}Ge_xC_y$ film, high repulsive interaction between Ge and C in the Si lattice plays a dominant role in determining the composition profiles of the film, and attractive interaction between Si and C affects the final film morphology. Figure 5.17 shows a STM image of a surface of a $Si_{0.478}Ge_{0.478}C_{0.044}$ film with a thickness of 4 ML together with a close-up image of the boxed region. The surface exhibits the onset of three-dimensional (3D) islanding, indicating the local increase of the Ge fraction around the island where the film locally exceeds the critical thickness for islanding. It should be noted that these islands are different from those due to C-induced Ge islanding (Schmidt *et al.*, 1997; Leifeld *et al.*, 1999a), since they are not formed at the beginning of the growth as shown in Fig. 5.16(b). The close-up atomic image in Fig. 5.17 shows fairly flat terraces consisting of the $c(4 \times 4)$ reconstructed structure caused by

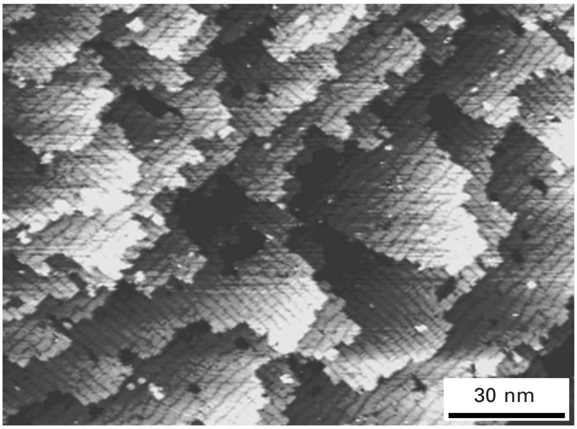


5.17 STM image of $Si_{0.478}Ge_{0.478}C_{0.044}$ films at 4 ML thickness grown on Si(001) substrates at 600°C. A magnified image of the boxed area is also shown, where examples of the $c(4 \times 4)$ and $(2 \times n)$ reconstruction are indicated.

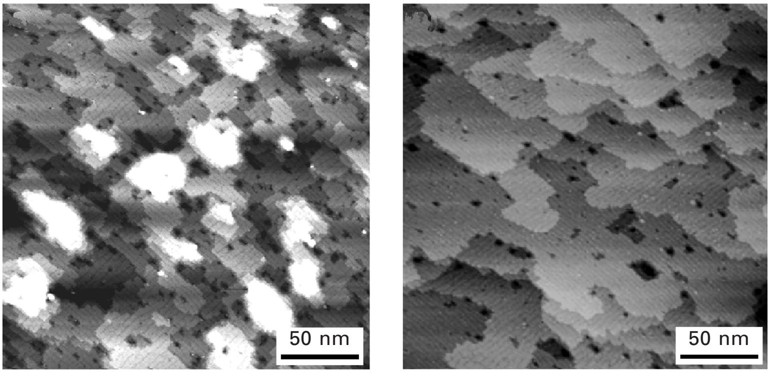
significant C condensation on the growing surface (Uhrberg *et al.*, 1992; Leifeld *et al.*, 1999b) and the $(2 \times n)$ reconstruction composed of buckled dimers and missing dimer rows. These results clearly show the formation of C-rich and Ge-rich regions in the film, which is presumably driven by the phase separation between Si–C and Si–Ge during growth when the film contains a high concentration of C.

Although the increased C atoms degrade the film morphology, the addition of a thin (1–2 ML) SiGe interlayer between the $Si_{1-x-y}Ge_xC_y$ film and the Si substrate drastically improves the film structure, leading to a planar morphology even with large C fractions present in the film (Sakai *et al.*, 2001b). Figure 5.18 is a STM image of a sample which has a structure of $Si_{0.473}Ge_{0.473}C_{0.054}$ (6 ML)/$Si_{0.5}Ge_{0.5}$ (1 ML)/Si(001). A planar, step and terrace surface morphology is formed even though the film contains a 4.6% average C fraction and no $c(4 \times 4)$ reconstructions are observable. In this case, the $Si_{1-x}Ge_x$ interlayer explicitly plays a role in suppressing C condensation and Si–C/Si–Ge phase separation during the film evolution.

Similarly to the above growth engineering, sequential alternate deposition, in which Si and Ge co-deposition and C deposition are performed alternately, has also been attempted in order to suppress the local phase separation, the 3D island growth, and the defect formation (Takeuchi *et al.*, 2005). Figure 5.19 shows a comparison of surface atomic morphologies between the 5-ML-thick $Si_{0.769}Ge_{0.183}C_{0.048}$ layers formed by the co-deposition of Si, Ge and C and by the alternate deposition of 1-ML-thick $Si_{0.793}Ge_{0.207}$ and 0.048-ML-thick C layer on Si(001). In the case of co-deposition shown in Fig. 5.19(a), a rough surface with small terraces and 3D islands is formed. This seems to result from the prevention of the conformal step flow growth due to the positive (normal) Ehrlich–Schwoebel barrier (Ehrlich and Hudda, 1966;



5.18 STM image of a sample with a $Si_{0.473}Ge_{0.473}C_{0.054}$ (6 ML)/ $Si_{0.5}Ge_{0.5}$(1 ML)/Si(001) structure. Note that a planar surface morphology is formed in spite of the large C fraction in the sample.



(a) (b)

5.19 STM images of 5-ML-thick $Si_{0.769}Ge_{0.183}C_{0.048}$ layers formed by (a) the co-deposition of Si, Ge and C, and (b) the alternate deposition of 1-ML-thick $Si_{0.793}Ge_{0.207}$ and 0.048-ML-thick C layers. In the co-deposition, a rough surface consisting of small terraces, defects as dark spots, and 3D islands with a height of approximately 2 nm is observed on the surface, while no 3D islands with reduced defects are seen in the alternate deposition.

Schwoebel and Shipsey, 1966) of defects seen as dark spots on the surface. On the other hand, in Fig. 5.19(b), no 3D islands are seen and the aligned steps are still formed. The density of defects is relatively low compared with the co-deposition case.

In both the interlayer addition and the alternate deposition processes, C atom migration on the growing surface is effectively restrained because the

Ge atoms, which give rise to the repulsive interaction force to the deposited C atoms, are uniformly distributed on the $Si_{1-x}Ge_x$ surface. This effect leads to the suppression of defect formation in the initial stages of $Si_{1-x-y}Ge_xC_y$ growth and consequently increases the critical thickness up to which layer-by-layer growth of $Si_{1-x-y}Ge_xC_y$ proceeds.

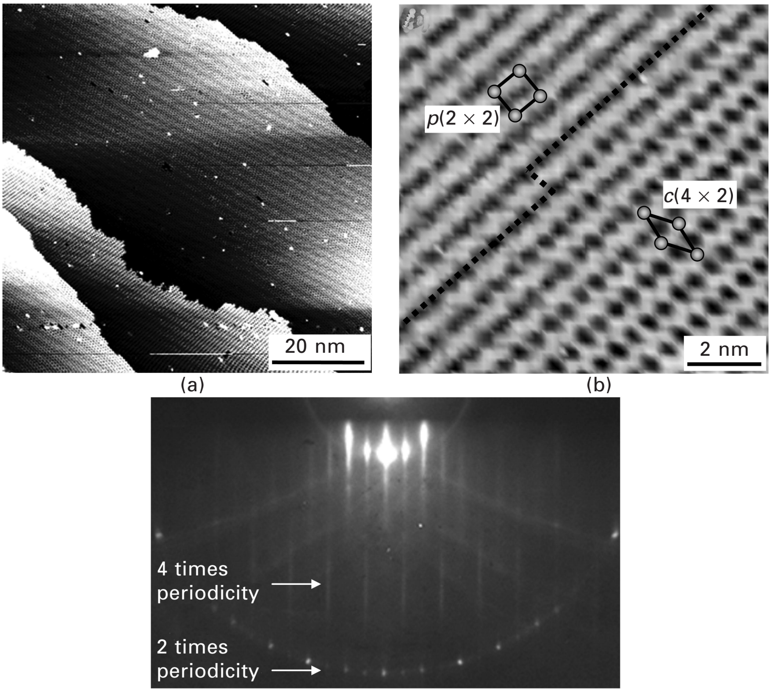
Technological application of epitaxial $Si_{1-x-y}Ge_xC_y$ and $Si_{1-x}C_x$ films is now proceeding mainly in the fabrication of high performance Si-based heterobipolar transistors and tensile Si channel MOSFETs with source-drain stressors. Understanding the fundamental mechanism of $Si_{1-x-y}Ge_xC_y$ alloy film evolution on an atomic scale has only just begun and the introduction of a high concentration of C into SiGe is still a challenging issue.

5.3.5 Oxygen reaction with clean Ge surfaces

In spite of the high potentiality of Ge, process technologies required for the fabrication of Ge-based MOSFETs are still immature. Especially, contrary to the well-known SiO_2/Si system, the quality and stability of germanium oxide is a principal issue when applied to the device. This yields technical difficulties in surface passivation and insulator formation on Ge when used as a channel material. In general, interaction between the solid surface and the vapor is highly sensitive to the ambient pressure and temperature. For the Si–oxygen system, Lander and Morrison (1962) carried out pioneering work that actually offered a significant insight into the categorization of the Si–oxygen chemical reaction. The occurrence of either oxidation or etching on the Si surface critically depends on the oxygen partial pressure and the substrate temperature. In this subsection, we focus on the Ge–oxygen system. Using the MBE chamber, a clean surface of Ge can be prepared under UHV conditions and then the surface is exposed to an oxygen gas with a controlled partial pressure and substrate temperature. This simple process realizes the Ge–oxygen chemical reaction which can be observed by in-situ STM. Formation of several types of nanostructures on the clean Ge surface due to the oxidation and etching processes is characterized on an atomic scale.

Figure 5.20(a) is a STM image of a clean Ge(001) surface consisting of step and terrace structures. An atomic-scale STM image of Fig. 5.20(b) clearly reveals a typical feature of a clean Ge surface which shows the coexistence of $p(2 \times 2)$ and $c(4 \times 2)$ reconstructed structures (Zandvliet, 2003), as is also observed in a RHEED pattern shown in Fig. 5.20(c). In order to expose such a Ge surface to oxygen, oxygen gas is introduced through a variable leak valve to the MBE chamber. After reaction at a certain oxygen pressure, substrate temperature and time, the sample is quenched to room temperature to perform STM observation.

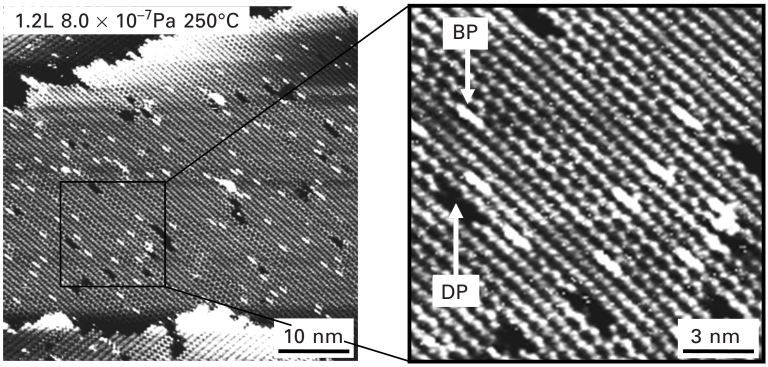
Figure 5.21 shows STM images of a Ge surface subjected to 1.2 L exposure


20 nm
p(2 × 2)
c(4 × 2)
2 nm
4 times periodicity
2 times periodicity

(a) (b)

(c)

5.20 STM images of a clean Ge surface showing (a) step and terrace structure and (b) a domain boundary between *p*(2 × 2) and *c*(4 × 2) atomistic structures on the terrace. (c) RHEED pattern of a clean Ge surface.


1.2L 8.0 × 10^{-7}Pa 250°C
10 nm
BP
DP
3 nm

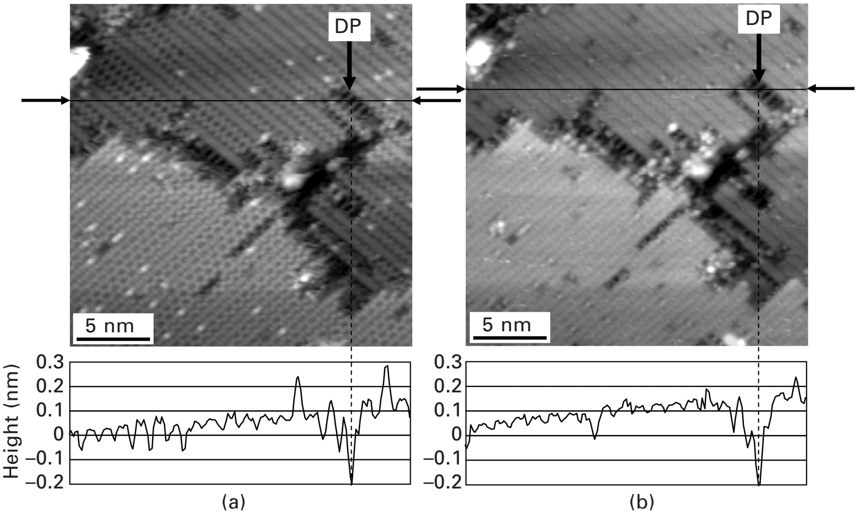
5.21 STM image of a clean Ge surface subjected to oxygen exposure of 1.2 L with an oxygen partial pressure of 8.0×10^{-7} Pa at a substrate temperature of 250°C. A magnified image of the boxed region is also shown, where BP and DP can be clearly identified.

with an oxygen partial pressure of 8.0×10^{-7} Pa at a substrate temperature of 250°C. The magnified image of the boxed region shows that two typical contrasts appear on the terrace region: bright products (BP) placed on and along the dimer row and dark products (DP) along the dimer row. BP likely originates from oxygen-related nuclei initially formed on the Ge surface (Fukuda and Ogino, 1997). As shown later in Fig. 5.23(c), BP becomes larger for a longer time exposure of oxygen, resulting in oxide clusters.

For DP, dual-bias STM images of the same area clarify its nature in more detail. Figures 5.22(a) and 5.22(b) show filled- and empty-state images, respectively, where DP exhibits the dimer rows aligned perpendicularly to the adjacent dimers. Cross-sectional profiles show that they are approximately 0.15 nm lower than the adjacent area. Thus it is clear that the DP are simply missing dimer rows, probably caused by the etching reaction of Ge and oxygen leading to volatile Ge mono-oxide formation as

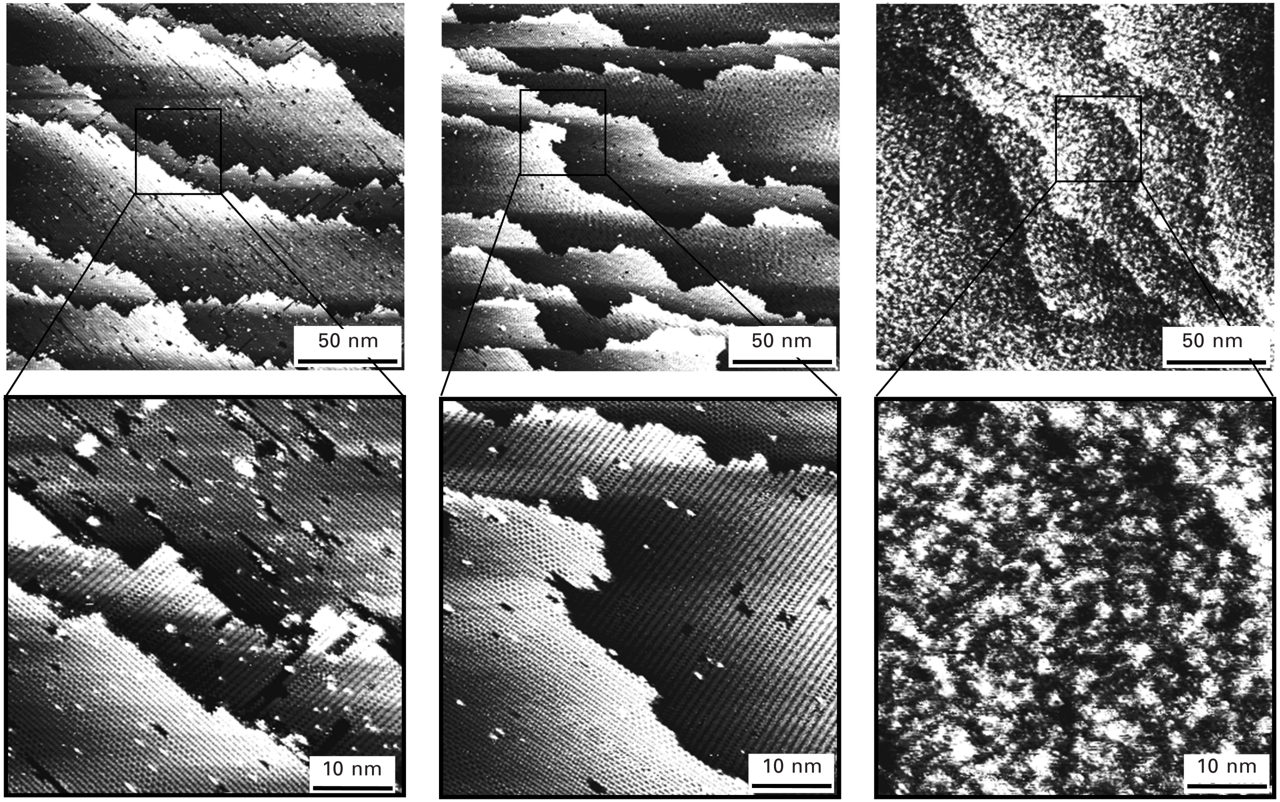
$$2\text{Ge(s)} + \text{O}_2\text{(g)} = 2\text{GeO(g)} \tag{5.3}$$

As shown in Fig. 5.23(a), with increasing the oxygen exposure, the surface exhibits DP elongated along the dimer row direction, keeping the width almost the same, which results from the local progressive etching reaction. Note that both steps parallel and perpendicular to the dimer row become



(a) (b)

5.22 Dual-bias STM images of the same area of the sample subjected to oxygen exposure of 6 L with an oxygen partial pressure of 8.0×10^{-7} Pa at a substrate temperature of 250°C: (a) filled-state image at a sample bias $V_b = -2.0$ V and (b) empty-state image at $V_b = 2.0$ V. Both images show that DP exhibits the dimer rows which are aligned perpendicularly to the adjacent dimers. Cross-sectional profiles show that they are approximately 0.15 nm lower than the adjacent area.



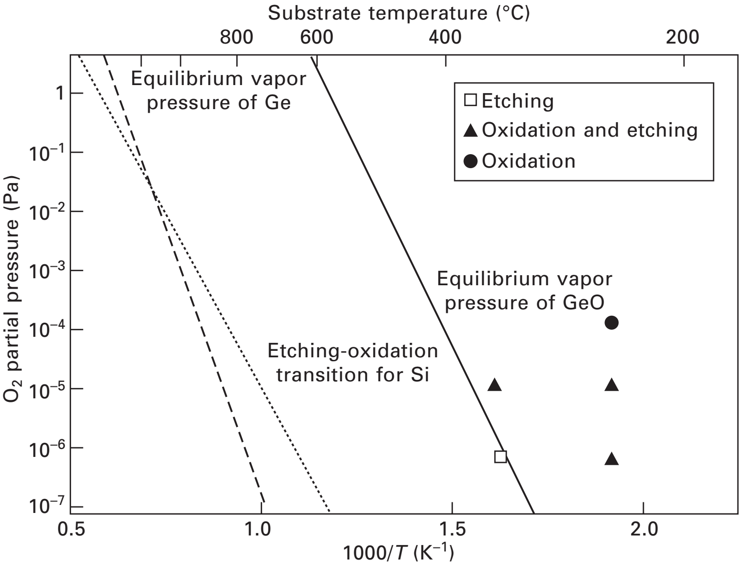
(a) 6L, OPP 8.0×10^{-7} Pa, 250°C (b) 6L, OPP 8.0×10^{-7} Pa, 350°C (c) 240 L, OPP 1.3×10^{-4} Pa, 250°C

5.23 STM images of a clean Ge surface subjected to oxygen exposure of 6 L with an oxygen partial pressure (OPP) of 8.0×10^{-7} Pa at a substrate temperature of (a) 250°C and (b) 350°C, and (c) 240 L with an OPP of 1.3×10^{-4} Pa at 250°C. Magnified images of the boxed regions are also shown.

straight due to the etching. Supposing that the step kink, i.e. the protruded Ge dimer, acts as a preferential reaction site for the oxygen and the reaction rate depends on the degree of step protrusion, the etching reaction proceeds in such a way that the step becomes flattened to produce the highly straight shape of both S_A and S_B steps.

Comparison between Figs 5.23(a) and 5.23(b) shows a surface structure variation depending on the substrate temperature under the same oxygen exposure. When increasing the temperature up to 350°C, straight steps are no longer formed but step meandering occurs, in which the steps are etched without showing any preferential directions and both DP and BP sharply decrease. On the other hand, when increasing the oxygen partial pressure at 250°C, the surface reaction mode changes from etching to oxidation. As shown in Fig. 5.23(c), a large number of oxide clusters cover the whole surface, keeping the initial step morphology, so that the oxidation reaction becomes dominant.

From the aforementioned results, the reaction mode dependence on the oxygen partial pressure and substrate temperature can be categorized as in Fig. 5.24. Higher substrate temperature and lower oxygen partial pressure lead



5.24 Oxygen reaction modes on clean Ge(001) surfaces depending on the oxygen partial pressure and the substrate temperature, identified by the present STM observations. The dotted line indicates the border between oxidation and etching reactions for Si surfaces obtained by Lander and Morrison (1962). The solid and broken lines indicate equilibrium vapor pressures of GeO (Jolly and Latimer, 1952) and Ge, respectively.

to the etching reaction forming a clean Ge surface, whereas lower substrate temperature and higher oxygen partial pressure lead to the oxidation reaction forming an oxidized Ge surface. Figure 5.24 also includes the etching–oxidation transition line for Si given by Lander and Morrison (1962) as well as the lines showing equilibrium vapor pressures of Ge and GeO (Jolly and Latimer, 1952). As expected, the etching–oxidation transition occurs at a much higher oxygen partial pressure and a lower substrate temperature than those for Si, and the transition of etching to oxidation approximately corresponds to the equilibrium vapor pressure of GeO. This result clearly demonstrates that the stability of germanium oxide can be guaranteed at a lower temperature and a higher oxygen partial pressure compared with the Si case.

5.4 Future trends

In this chapter, practical techniques of MBE and some example studies of MBE for SiGe and related materials have been reviewed. In this materials system, the growth behavior is principally governed by the lattice mismatch between Si and Ge, and thus strain introduction and relaxation are matters for investigation. Elucidation of the interplay between the morphology and the defect should be a key to controlling the formation of SiGe nanostructures which will be installed in future nanoelectronics. Since the nanostructures are strongly influenced by the defect properties, defect engineering is crucial.

Contrary to the complete solubility in the solid state of Si and Ge, the addition of Sn or C elements into Si or Ge drastically alters the growth behavior of thin films, in which the interaction of the atomic species determines the final film morphology. Although the formation of semiconductor alloys with such immiscible species is still a challenging issue, a quest for a novel intrinsic property of the alloy motivates the development of new growth schemes. An improvement from layer-by-layer to atom-by-atom growth control in the MBE technique surely overcomes this issue.

It is now accepted that MBE does not seem to be a suitable technique for mass production due to its relatively high workload of machine maintenance and low throughput of sample production compared with other growth techniques based on chemical vapor deposition. Nevertheless, the importance of MBE is not diminished. In particular, for fundamental research into state-of-the-art technologies of thin film growth and device fabrication, MBE is one of the most powerful tools for investigating the essential formation mechanisms of thin films and nanostructures as well as exploring their new functionalities. Deep understanding of such mechanisms possibly brings out the latent power of SiGe thin films and related nanostructures, which leads to novel sophisticated Si-based devices with high performance.

5.5 References

Alerhand O L, Vanderbilt D, Meade R D and Joannopoulos J D (1998), 'Spontaneous formation of stress domains on crystal surfaces', *Phys. Rev. Lett.*, 61, 1973–1976.

Berti M, De Salvador D, Drigo A V, Romanato F, Stangl J, Zerlauth S, Schäffler F and Bauer G (1998), 'Lattice parameter in $Si_{1-y}C_y$ epilayers: Deviation from Vegard's rule', *Appl. Phys. Lett.*, 72, 1602–1604.

Chadi D J (1987), 'Stabilities of single-layer and bilayer steps on Si(001) surfaces', *Phys. Rev. Lett.*, 59, 1691–1694.

Christian J W (1982), *Transformation in Metal and Alloys*, Oxford, Pergamon Press, p. 448.

Eaglesham D J and Cerullo M (1990), 'Dislocation-free Stranski–Krastanow growth of Ge on Si(100)', *Phys. Rev. Lett.*, 64, 1943–1946.

Ehrlich G and Hudda F (1966), 'Atomic view of surface self-diffusion: Tungsten on tungsten', *J. Chem. Phys.*, 44, 1039–1049.

Fang Y Y, Tolle J, Roucka R, Chizmeshya A V G, Kouvetakis J, D'Costa V R and Menéndez J (2007), 'Perfectly tetragonal, tensile-strained Ge on $Ge_{1-y}Sn_y$ buffered Si(100)', *Appl. Phys. Lett.*, 90, 061915-1–3.

Fischetti M V and Laux S E (1996), 'Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys', *J. Appl. Phys.*, 80, 2234–2252.

Foster M, Darlington B, Scharff J and Campion A (1997), 'Surface chemistry of alkylsilanes on Si(100)2 × 1', *Surf. Sci.*, 375, 35–44.

Fukuda T and Ogino T (1997), 'Initial oxygen reaction on Ge(100) 2 × 1 surfaces', *Phys. Rev. B*, 56, 13190–13193.

Gu S, Wang R, Zhang R and Zheng Y (1996), 'Simulation model to very low pressure chemical vapor deposition of SiGe alloy', *J. Vac. Sci. Technol. A*, 14, 3256–3260.

Hammer M, LeGoues F K, Telsoff J, Reuter M C and Tromp R M (1996), 'In situ ultrahigh vacuum transmission electron microscopy studies of hetero-epitaxial growth I. Si(001)/Ge', *Surf. Sci.*, 349, 129–144.

Harris J J, Joyce B A and Dobson P J (1981), 'Oscillations in the surface structure of Sn-doped GaAs during growth by MBE', *Surf. Sci.*, 103, L90–L96.

Hirayama H, Tatsumi T, Ogura A and Aizaki N (1987), 'Gas source silicon molecular beam epitaxy using silane', *Appl. Phys. Lett.*, 51, 2213–2215.

Ishikawa Y, Wada K, Cannon D D, Liu J, Luan H C and Kimerling L C (2003), 'Strain-induced band gap shrinkage in Ge grown on Si substrate', *Appl. Phys. Lett.*, 82, 2044–2046.

Iwasaki F, Tomitori M and Nioshikawa O (1991), 'STM study of epitaxial growth of Ge on Si(001)', *Surf. Sci. Lett.* 253, L411–L416.

Iyer S S, Eberl K, Goorsky M S, LeGoues F K, Tsang J C and Cardone F (1992), 'Synthesis of $Si_{1-y}C_y$ alloys by molecular beam epitaxy', *Appl. Phys. Lett.*, 60, 356–358.

Jolly W L and Latimer W M (1952), 'The equilibrium Ge(s) + GeO_2(s) = 2GeO(g). The heat of formation of germanic oxide', *J. Am. Chem. Soc.*, 74, 5757–5758.

Kelires P C (1997), 'Short-range order, bulk moduli, and physical trends in c-$Si_{1-x}C_x$ alloys', *Phys. Rev. B*, 55, 8784–8787.

Kim H, Taylor N, Bramblett T R and Greene J E (1998), 'Kinetics of $Si_{1-x}Ge_x$(001) growth on Si(001)2 × 1 by gas-source molecular-beam epitaxy from Si_2H_6 and Ge_2H_6', *J. Appl. Phys.*, 84, 6372–6381.

Lam A M, Zheng Y J and Engstrom J R (1997), 'Gas-surface reactivity in mixed-crystal systems: the reaction of GeH_4 and Ge_2H_6 on Si surfaces', *Surf. Sci.*, 393, 205–221.

Lander J J and Morrison J (1962), 'Low voltage electron diffraction study of the oxidation and reduction of silicon', *J. Appl. Phys.*, 33, 2089–2092.

LeGoues F K, Copel M and Tromp R M (1989), 'Novel strain-induced defect in thin molecular-beam epitaxy layers', *Phys. Rev. Lett.*, 63, 1826–1829.

Leifeld O, Beyer A, Muller E, Kern K and Grutzmacher D (1999a), 'Formation and ordering effects of C-induced Ge dots grown on Si (001) by molecular beam epitaxy', *Mat. Sci. Eng. B.*, 74, 222–226.

Leifeld O, Grutzmacher D, Muller B, Kern K, Kaxiras E and Kelires P C (1999b), 'Dimer pairing on the C-alloyed Si(001) surface', *Phys. Rev. Lett.*, 82, 972–975.

Madelung O (1996), *Semiconductors – Basic Data*, 2nd revised edn, Berlin, Springer, p. 22.

Medeiros-Ribeiro G, Bratkovski A M, Kamins T I, Ohlberg D A A and Williams R S (1998), 'Shape transition of germanium nanocrystals on a silicon (001) surface from pyramids to domes', *Science*, 279, 353–355.

Men F K, Packard W E and Webb M B (1988), 'The Si(100) surface under an externally applied stress', *Phys. Rev. Lett.*, 61, 2469–2471.

Mo Y W, Savage D E, Swartzentruber B S and Lagally M G (1990), 'Kinetic pathway in Stranski–Krastanov growth of Ge on Si(001)', *Phys. Rev. Lett.*, 65, 1020–1023.

Mochizuki S, Sakai A, Taoka N, Nakatsuka O, Takeda S, Kimura S, Ogawa M and Zaima S (2006), 'Local strain in SiGe/Si heterostructures analyzed by X-ray microdiffraction', *Thin Solid Films* 508, 128–131.

Murata T and Suemitsu M (2004), 'GeH_4 adsorption on Si(001) at RT: transfer of H atoms to Si sites and atomic exchange between Si and Ge', *Appl. Surf. Sci.*, 224, 179–182.

Neave J H, Joyce B A, Dobson P J and Norton N (1983), 'Dynamics of film growth of GaAs by MBE from RHEED observations', *Appl. Phys. A*, 31, 1–8.

Ni W X, Henry A, Larsson M I, Joeisson K and Hansson G V (1994), 'High quality Si/$Si_{1-x}Ge_x$ layered structures grown using a mass-spectrometry controlled electron-beam evaporation system', *Appl. Phys. Lett.*, 65, 1772–1774.

Oda K, Ohue E, Suzumura I, Hayami R, Kodama A, Simamoto H and Washio K (2003), 'High-performance self-aligned SiGeC HBT with selectively grown $Si_{1-x-y}Ge_xC_y$ base by UHV/CVD', *IEEE Trans. Electron. Devices*, 50, 2213–2220.

Okinaka M, Hamana Y, Tokuda T, Ohta J and Nunoshita M (2003), 'MBE growth mode and C incorporation of GeC epilayers on Si(001) substrates using an arc plasma gun as a novel C source', *J. Cryst. Growth*, 249, 78–86.

Osten H J, Bugiel E and Zaumseil P (1994), 'Growth of an inverse tetragonal distorted SiGe layer on Si(001) by adding small amounts of carbon', *Appl. Phys. Lett.*, 64, 3440–3442.

Robbins D J, Glasper J L, Cullis A G and Leong W L (1991), 'A model for heterogeneous growth of $Si_{1-x}Ge_x$ films from hydrides', *J. Appl. Phys.*, 69, 3729–3732.

Ross F M, Tersoff J and Tromp R M (1998), 'Coarsening of self-assembled Ge quantum dots on Si(001)', *Phys. Rev. Lett.*, 80, 984–987.

Rücker H, Heinemann B, Bolze K D, Knoll D, Krüger D, Kurps R, Osten H J, Schley P, Tillack B and Zaumseil P (1999), 'Dopant diffusion in C-doped Si and SiGe: physical model and experimental verification', *Tech. Dig. Int. Electron. Device Meet.*, p. 345.

Sakai A and Tatsumi T (1993), 'Defect-mediated island formation in Stranski–Krastanov growth of Ge on Si(001)', *Phys. Rev. Lett.*, 71, 4007–4010.

Sakai A and Tatsumi T (1994), 'Defect and island formation in Stranski–Krastanov growth of Ge on Si(001)', *Mater. Res. Soc. Symp. Proc.*, 137, 343–346.

Sakai A, Sugimoto K, Yamamoto T, Okada M, Ikeda H, Yasuda Y and Zaima S (2001a), 'Reduction of threading dislocation density in SiGe layers on Si (001) using a two-step strain-relaxation procedure', *Appl. Phys. Lett.*, 79, 3398–3400.

Sakai A, Torige Y, Okada M, Ikeda H, Yasuda Y and Zaima S (2001b), 'Atomistic evolution of $Si_{1-x-y}Ge_xC_y$ thin films on Si(001) surfaces', *Appl. Phys. Lett.*, 79, 3242–3244.

Sakai A, Taoka N, Nakatsuka O, Zaima S and Yasuda Y (2005), 'Pure-edge dislocation network for strain-relaxed SiGe/Si(001) systems', *Appl. Phys. Lett.*, 86, 221916-1–3.

Sakamoto K, Sakamoto T, Nagao S, Hashigutchi G, Kuniyoshi K and Bando Y (1987), 'Reflection high-energy electron diffraction intensity oscillations during Ge_xSi_{1-x} MBE growth on Si(001) substrates', *Jpn. J. Appl. Phys.*, 26, 666–670.

Schmidt O G, Lange C, Eberl K, Kienzle O and Ernst F (1997), 'Formation of carbon-induced germanium dots', *Appl. Phys. Lett.*, 71, 2340–2342.

Schwoebel R L and Shipsey E J (1966), 'Step motion on crystal surfaces', *J. Appl. Phys.*, 37, 3682–3686.

Suemitsu M, Hirose F, Takakuwa Y and Miyamoto N (1990), 'Growth kinetics in silane gas-source molecular beam epitaxy', *J. Cryst. Growth*, 105, 203–208.

Takeda S, Kimura S, Sakata O and Sakai A (2006), 'Development of high-angular-resolution microdiffraction system for reciprocal space map measurements', *Jpn. J. Appl. Phys.*, 45, L1054–L1056.

Takeuchi S, Nakatsuka O, Wakazono Y, Sakai A, Zaima S and Yasuda Y (2005), 'Initial growth behaviors of SiGeC in SiGe and C alternate deposition', *Mat. Sci. Semicon. Process.*, 8, 5–9.

Takeuchi S, Sakai A, Yamamoto K, Nakatsuka O, Ogawa M and Zaima S (2007), 'Growth and structure evaluation of strain-relaxed $Ge_{1-x}Sn_x$ buffer layers grown on various types of substrates', *Semicond. Sci. Technol.*, 22, S231–S235.

Takeuchi S, Shimura Y, Nakatsuka O, Zaima S, Ogawa M and Sakai A (2008a), 'Growth of highly strain-relaxed $Ge_{1-x}Sn_x$/virtual Ge by a Sn precipitation controlled compositionally step-graded method', *Appl. Phys. Lett.*, 92, 231916-1–3.

Takeuchi S, Sakai A, Nakatsuka O, Ogawa M and Zaima S (2008b), 'Tensile strained Ge layers on strain-relaxed $Ge_{1-x}Sn_x$/virtual Ge substrates', *Thin Solid Films*, 517, 159–162.

Uhrberg R I G, Northrup J E, Biegelesen D K, Bringans R D and Swartz L E (1992), 'Atomic structure of the metastable $c(4 \times 4)$ reconstruction of Si(100)', *Phys. Rev. B*, 46, 10251–10256.

Zandvliet H J W (2003), 'The Ge(001) surface', *Phys. Rep.*, 388, 1–40.

6

Silicon–germanium (SiGe) crystal growth using chemical vapor deposition

B. TILLACK, IHP and Technische Universität Berlin, Germany
and J. MUROTA, Tohoku University, Japan

Abstract: In this chapter the main aspects of CVD tools and CVD heteroepitaxy processing, including Si surface treatment before epitaxy, are described. Low-temperature SiGe growth is discussed, especially by reviewing data regarding the growth rate and the Ge fraction as functions of different process parameters (partial pressures, temperatures). For the description of the growth mechanism a Langmuir-type adsorption scheme is shown. *In-situ* doping of SiGe by B and P and its usage for HBTs are reviewed, followed by a discussion of the atomic layer processing approach for doping. Finally an outlook is given summarizing topics for further development and improvement.

Key words: SiGe, heteroepitaxy, CVD epitaxy, B and P doping, atomic layer doping.

6.1 Introduction

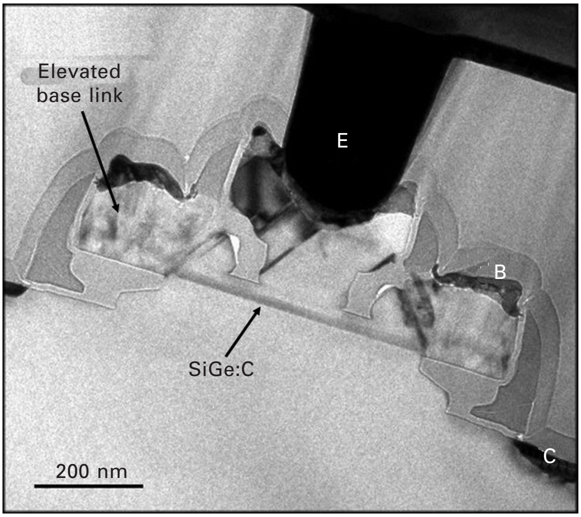
Using silicon–germanium on Si to create heterostructures with new properties has been intensively investigated over the last decades. Numerous research results have been published, demonstrating the capability of the material for applications in a wide range of electronic devices. The success story of SiGe is very much related to the development of SiGe heterojunction bipolar transistor technology (SiGe HBT), which makes use of SiGe bandgap engineering and strain (1–3). The HBT technology has reached an industrial level with increasing market potential for very high frequencies such as high-speed communication, radar application, mm-wave THz imaging and sensing (4).

Now SiGe and Si strained layers are becoming increasingly important for advanced CMOS technologies, for instance for higher mobility channel material (5, 6). The first report on growth of epitaxial SiGe layers dates back to 1962 (7), and especially the pioneering works of Kasper and Herzog (see e.g. 8) and Meyerson (see e.g. 9) were important milestones in the development of SiGe strained layer epitaxy. Starting with molecular beam epitaxy (MBE), the main material properties and heteroepitaxy characteristics were investigated. A main milestone in the industrialization of the SiGe technology was the demonstration of the stability and high volume production capability of SiGe strained layer epitaxy for HBT base deposition using chemical vapor deposition (CVD). First, ultra high vacuum CVD (UHVCVD) (10) was used because very low moisture and oxygen levels are required for the low-temperature epitaxy. But it is possible to fulfill the critical requirements regarding low impurity levels using a deposition system that works over the whole pressure range, from UHV through low pressure CVD (LPCVD) and reduced pressure CVD (RPCVD) (11–15) to atmospheric pressure CVD (APCVD) (16–21). Especially, the demonstration of the ability to grow device-quality SiGe layers without using UHV deposition techniques has greatly influenced the development of strained layer deposition process technology and tools.

Another important step in the success story of strained layer epitaxy for HBT application was the incorporation of carbon into SiGe (22, 23). It was shown that C can significantly suppress B diffusion without negative impact on device parameters (23, 24). The first BiCMOS technology using SiGe:C HBTs was demonstrated by IHP in 1999 (25).

Today SiGe:C strained layer epitaxy is widely used in state-of-the-art HBT technologies. As an example Fig. 6.1 shows a TEM cross-section of an HBT which is integrated in a 0.13 μm BiCMOS technology (26).

In this chapter we will review basic aspects of CVD SiGe epitaxy, considering also the different CVD techniques as the deposition process. A main part of the discussion is focused on doping of SiGe. Here as well as standard *in-situ* doping of SiGe as atomic layer doping is discussed.



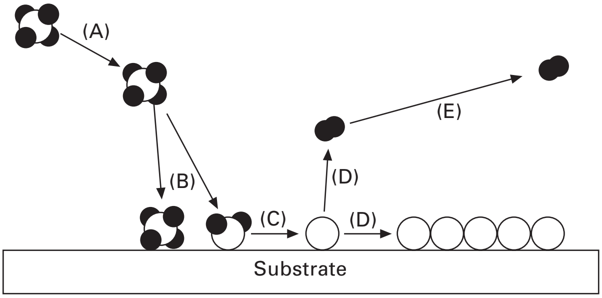
6.1 TEM cross-section of an HBT integrated in BiCMOS technology (26).

6.2 Epitaxial growth techniques – chemical vapor deposition (CVD) (ultra high vacuum CVD (UHVCVD), low pressure CVD (LPCVD), atmospheric pressure CVD (APCVD), plasma enhanced CVD (PECVD))

The epitaxial deposition process from the gas phase is the result of phenomena in the gas phase and at the surface of the substrate. The main steps of the deposition are (Fig. 6.2):

(A) Transport of the precursors with the carrier gas to the reaction chamber and diffusion to the wafer surface
(B) Adsorption or chemisorption of the precursors at the wafer surface
(C) Surface reaction
(D) Layer growth, desorption and diffusion of volatile reaction products
(E) Transport of reaction products with the carrier gas out of the reaction chamber.

Classical epitaxy processes, e.g. for a thick epitaxial Si layer on Si as starting wafer material for CMOS, use a high deposition temperature (>1000°C). Strained layer epitaxy of SiGe requires a low deposition temperature (typically 500–700°C) to prevent relaxation and three-dimensional growth and to control the deposition rates for thin epitaxial layers. At low temperature the deposition process is much more sensitive to contamination on the wafer surface compared to deposition at high temperature. Therefore it is crucial to keep the contamination level, especially moisture and oxygen, as low as possible on the wafer before loading into the deposition tool and to guarantee a low contamination level during handling and deposition. Based on the chemical equilibrium data for the Si/O_2/H_2O/SiO_2 system reported by Smith and Ghidini (27, 28), Meyerson *et al.* (29, 30) developed a hot-wall ultra-high vacuum CVD system in order to avoid the oxidation of the Si surface. Low-temperature epitaxial growth of Si using a SiH_4–H_2 gas mixture was



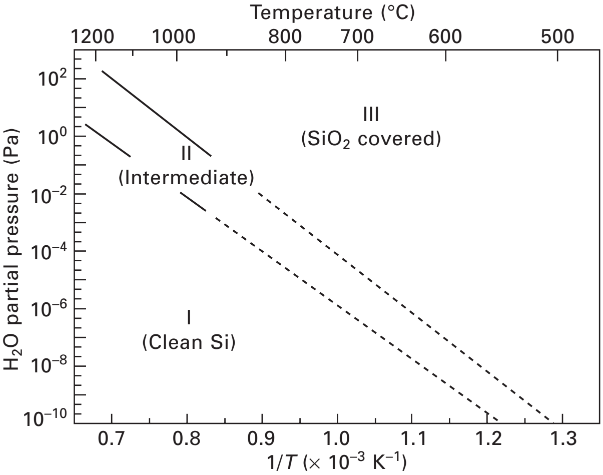
6.2 Main steps of a CVD process: for (A)–(E) see text.

obtained. A base pressure in the range of 10^{-7} Pa and a deposition pressure of about 0.1 Pa were used in order to achieve an oxide-free surface at a temperature below 800°C, extrapolating the high-temperature data (28) for the formation of oxide on the Si(100) surface as shown in Fig. 6.3, and assuming the existence of a 1 ppm background contamination level.

Murota *et al.* (31) developed an ultra-clean hot-wall LPCVD system using a deposition pressure of around 10 Pa. Based on optimized reactor design resulting in a very low contamination level in the reactor and improved quality of the used gases, low-temperature Si-selective epitaxial growth at 650–850°C using a SiH_4–H_2 gas mixture was demonstrated. Sedgwick *et al.* (32) developed an ultra-clean APCVD process with H_2 point-of-use purifier and realized low-temperature selective epitaxial growth at 600–850°C using a SiH_2Cl_2–H_2 gas mixture.

There were a couple of technical improvements that essentially boosted the development of low-temperature epitaxy from an interesting experimental process to an established manufacturing technique. The main milestones were:

- Improved reactor design resulting in low contamination level in the deposition system
- Usage of load look to reduce moisture contamination level
- Precursor gases with low contamination level
- Point-of-use purification of carrier gases
- Improved temperature uniformity (e.g. better temperature control, wafer rotation) resulting in very good thickness uniformity of the epitaxial layers



6.3 Chemical equilibrium data for $Si/H_2O/SiO_2$ system (27, 28). The dashed lines show the extrapolation from the experimental data.

- In-situ cleaning to reduce deposition chamber coating and memory effects.

In Table 6.1 the different CVD techniques used for epitaxy are summarized.

There are hot-wall and cold-wall types of Si substrate heating methods. In the hot-wall type, the reactor is a cylindrical quartz tube inserted in a furnace. Since temperature control is easy for many wafers, high throughput and good uniformity of growth can be obtained. The disadvantages are the limitation of deposition pressure and the slow temperature response compared with cold-wall heating. In the latter, the wafers are generally set on a SiC-coated graphite susceptor heated by RF power or IR lamp irradiation outside the reactor cooled by air or water. Gibbons *et al.* (33) developed limited reaction processing (LRP) using rapid thermal annealing (RTA) in which the wafer is heated directly through a quartz window by IR lamps. In this method, the deposition process is controlled by ramping the temperature up and down, with the reactant gases flowing continuously. This rapid thermal CVD (RTCVD) technique is applied using deposition pressures ranging from ultra-high vacuum to atmospheric pressure. Early reports showed the capability of the RTCVD technique for low-temperature epitaxy using dichlorsilane (DCS) (14), silane (34), and disilane (35).

Most epitaxial growth techniques use thermal activation of the process. But especially for high growth rates at very low temperature, the use of plasma activated processing (PECVD) has been reported (36, 37).

Although low-temperature epitaxy is now an established manufacturing process (38), there is demand for further development from different applications. Two examples are lowering the process temperature (mainly the bake temperature) to improve the integration capability of the process (e.g. for advanced transistor engineering in CMOS), and optimization of the process regarding cost of ownership. Because throughput is one of the most important cost aspects, the development of batch epitaxy tools is again under consideration (39).

Table 6.1 Deposition and base pressure of different CVD techniques used for low-temperature Si and SiGe epitaxy

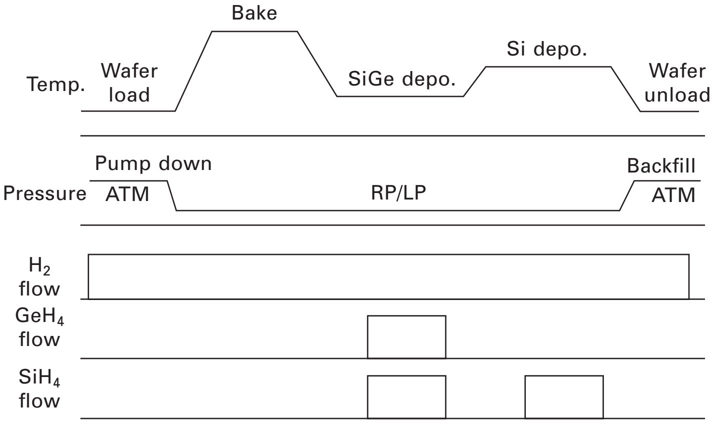
Technique	Deposition pressure (Pa)	Base pressure (Pa)	Activation	Reactor	Ref.
UHV	~0.1	$\sim10^{-7}$	Thermal	Batch/single	29, 30
LPCVD	10	$\sim10^{-6}$	Thermal	Batch	31
LPCVD	50	3	Thermal	Batch	39
LPCVD	~200	10^{-5}	Thermal	Single	10, 12
RPCVD	$\sim10^3$–10^4	~2	Thermal	Single	37, 38
APCVD	Atmospheric		Thermal	Single	16–21

6.3 Silicon–germanium (SiGe) heteroepitaxy by chemical vapor deposition (CVD)

SiGe heteroepitaxy on Si results in a strained layer or in partially or fully relaxed layers. If the thickness of the layer is smaller than the critical thickness (40–42), the layer grows pseudomorphically. The critical thickness depends on the total strain energy in the layer, which is mainly determined by the Ge fraction and the layer thickness. If a critical value for the strain in the layer is exceeded the film starts to relax, forming misfit dislocations (40, 43). In the case of thin SiGe layers used, e.g., for SiGe HBTs or in source drain areas of MOSFETs, the dislocations can cause pipes, degrading the electrical performance of the device (11). For these applications the CVD heteroepitaxy has to be optimized for the deposition of strained layers, preventing any relaxation. This affects the process conditions used. To prevent relaxation the deposition of thin layers has to be controlled in a proper way (usually by low deposition temperature and low growth rates).

Relaxed SiGe layers on Si are interesting as virtual substrates, e.g. for lattice engineering for Ge or III/V deposition (44). In this case high deposition rates and a higher temperature are desired, with the limitation to keep the growth in a 2D mode.

In Fig. 6.4 a typical process sequence for the deposition of SiGe capped with Si is described. After loading the Si wafer into the process chamber an H_2 bake step is performed to clean the Si surface directly before the deposition. For the bake high temperatures (typically 900–1000°C) are required. After the bake the temperature and the source gas flow (e.g. SiH_4, GeH_4) for the SiGe deposition are adjusted. Hydrogen is used as the carrier gas. For the Si cap layer deposition the temperature is increased. SiH_4 is used as precursor



6.4 Typical process sequence for epitaxial deposition of SiGe with Si cap.

for the Si growth. After the Si cap deposition the wafer is cooled down and unloaded. In the following section the Si surface preparation by ex-situ cleaning and *in-situ* H_2 bake, and the low-temperature Si and SiGe growth, are discussed in more detail.

6.3.1 Silicon surface preparation

The crystalline perfection of the deposited epitaxial layer is essentially determined by the perfection of the initial Si surface before the deposition. To create a Si surface with low contamination level and free of oxide, a combination of ex-situ wet chemical cleaning and *in-situ* H_2 bake in the reaction chamber before epitaxial deposition is applied (45, 46). An overview of the wet chemical cleaning procedures and their effect on contamination removal is summarized in Table 6.2 (47). Here two possible options for wet chemical cleaning are discussed, resulting in different surface passivation after wet chemical treatment. By applying SC-1/DHF dip/SC-2 (SC-1 = Standard Cleaning 1, DHF = diluted HF, SC-2 = Standard Cleaning 2; see Table 6.2), cleaning and stable passivation of the surface are obtained. During the final SC-2 treatment under standard conditions ($HCl + H_2O_2 + H_2O$ at 75–85°C) a thin, chemically grown oxide is formed, which covers the surface very well. This oxide has to be removed during the H_2 prebake. There is no issue if high bake temperatures could be applied, e.g. if the SiGe epitaxy is performed on blanket wafer and not as an integrated process step in CMOS or BiCMOS. To reduce the required bake temperature a modified cleaning sequence was applied, which uses a reduced temperature during the final SC-2 cleaning treatment (45). A lower temperature results in a thinner, chemically grown oxide. In this way the bake temperature could be reduced. The thickness of the oxide grown at a lower temperature during SC-2 is increased if the surface is exposed to air. Therefore there is the effect of a time delay between wet chemical cleaning and starting the epitaxy process on the bake temperature necessary to obtain an oxide-free Si surface.

Table 6.2 Contaminant removal and chemistry of standard cleaning procedures

Cleaning step*	Contamination removed	Chemistry	Mixing ratio	Temperature (°C)
SC-1	Organics, particles	$NH_4OH + H_2O_2 + H_2O$	1:1:5 to 0.25:1:5	50–85
SC-2	Metallic ions	$HCl + H_2O_2 + H_2O$	1:1:6 to 1:1:5	75–85
DHF	Metallic ions, native oxide	$HF + H_2O$	50:1 to 100:1	20–25

*SC-1 standard cleaning 1, SC-2 standard cleaning 2, DHF diluted HF.
Source: Kern and Puotinen (47).

The commonly used final wet cleaning treatment is diluted HF (DHF) dip followed by a deionized (DI) water rinse (48). In this case the surface is passivated by hydrogen generated during the final DHF dip. The passivation, in other words the hydrogen termination, is obtained using the conventional diluted HF etching of SiO_2 (49, 50), but it is not perfect due to the final rinse of DI water. Therefore, as for the SC-2 wet chemical cleaning described above, there is the effect of a time delay between the wet chemical cleaning and the start of the epitaxy process. If a critical delay time is exceeded, the hydrogen termination is resolved, resulting in increased oxygen contamination at the surface. Applying the bake temperature optimized for perfect hydrogen termination results in a defective surface and in this way in defective epitaxy layers.

After loading the ex-situ cleaned wafers into the reactor the H_2 prebake is performed. The oxide covering the surface is removed according to the following reaction scheme:

$$SiO_2 + Si \rightarrow 2SiO \text{ (volatile)} \qquad 6.1$$

$$SiO_2 + H_2 \rightarrow SiO \text{ (volatile)} + H_2O \qquad 6.2$$

The required bake temperature depends on the kind of oxide, the oxide coverage and the moisture level in the reaction chamber. Applying an optimized wet cleaning procedure (DHF dip last) and using a deposition tool with low moisture level in the reactor, the bake temperature can be reduced to below 800°C.

The bake temperature mainly determines the thermal budget of the epitaxy process itself. For the integration of epitaxy processes into CMOS or BiCMOS technologies the impact of the thermal budget on existing structures has to be considered. In this respect low-temperature *in-situ* cleaning procedures are highly desired. There are some encouraging experimental results using, e.g., Cl etching of the surface at low temperature (51), but there is no process available at manufacturing level.

6.3.2 Low-temperature SiGe growth

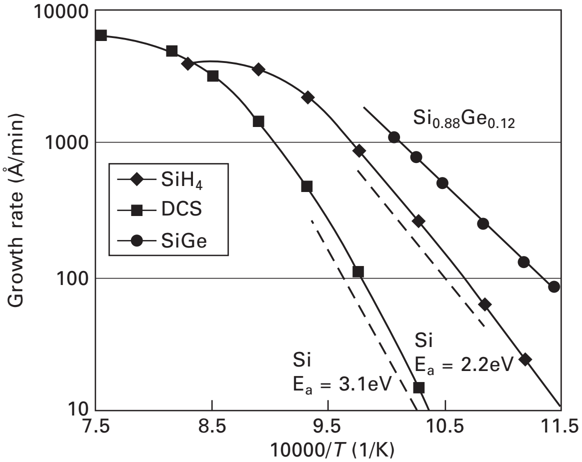
Today heteroepitaxy by CVD is in production with proven stability and manufacturability. In the following section key aspects and requirements of Si and SiGe CVD epitaxy will be discussed.

Considering the temperature dependence of Si epitaxy, there are two growth regimes. At high temperature the deposition rate is weakly dependent on the temperature only. The deposition is mass-transport limited. Activation energies in the order of 4–6 kcal/mol have been reported for different process conditions (52, 53).

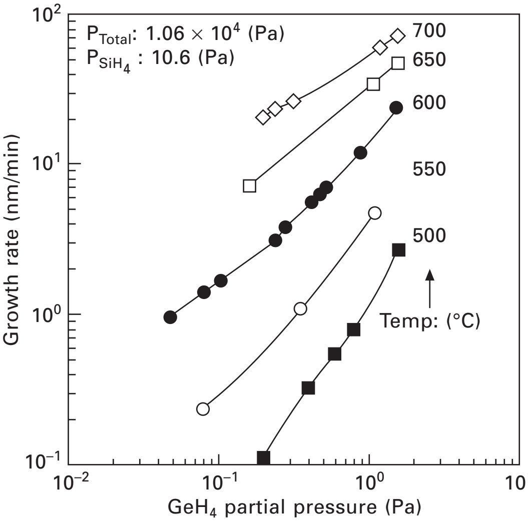
By lowering the deposition temperature the mechanism changes. The growth rate decreases exponentially with increasing reverse absolute temperature.

The deposition process is kinetically controlled. The elementary process determining the deposition process at low temperature is hydrogen desorption from the surface. Hydrogen has to be desorbed to allow the incorporation of Si atoms into the growing layer. Activation energies in the order of 42–50 kcal/mol have been reported depending on the chemistry and deposition systems used (54–63). This is in good agreement with the activation energy found for the desorption of hydrogen from Si(100) surfaces (64). In Fig. 6.5 the growth rates for Si and SiGe as functions of the reverse temperature are shown (38). For Si epitaxy the mass-transport-limited regime is visible at high temperature. In the low temperature region the process is kinetically controlled and is strongly dependent on temperature. Deposition of SiGe is much faster than that of Si under the same experimental conditions due to the catalytic effect of Ge on desorption of hydrogen.

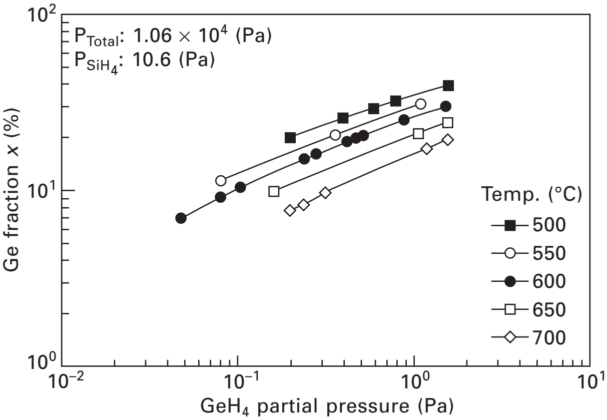
Tillack and Zaumseil (65) reported the growth rate as a function of the reciprocal temperature for different GeH_4 partial pressures. GeH_4 and SiH_4 are the sources for the SiGe layer growth. The activation energies obtained decrease with increasing GeH_4 partial pressure (increasing Ge fraction). For the lowest GeH_4 partial pressure the activation energy is about 44 kcal/mol, which is close to the value obtained for pure Si deposition. Adding GeH_4 enhances desorption of hydrogen, resulting in lower activation energies. Figure 6.6 shows the SiGe growth rate as a function of the GeH_4 partial pressure for different temperatures. At a given temperature the growth rate increases with increasing GeH_4 partial pressure. As discussed already for the impact of adding GeH_4 on the activation energy, the growth rate enhancement effect is due to lowering the hydrogen coverage at the surface by Ge. In Fig. 6.7



6.5 Epitaxial growth rate as a function of the reverse temperature (Arrhenius plot) for Si (silane and dichlorsilane) and SiGe, from (38).



6.6 SiGe growth rate as a function of the GeH_4 partial pressure at different deposition temperatures.



6.7 Ge fraction as a function of the GeH_4 partial pressure at different deposition temperatures.

the Ge fraction in the SiGe layer is shown as a function of the GeH_4 partial pressure for constant SiH_4 partial pressure and at different temperatures. The Ge fraction increases with increasing GeH_4 partial pressure for a given temperature. The incorporation of Ge is affected by the deposition temperature. At constant GeH_4 and SiH_4 partial pressures, the Ge fraction decreases with increasing temperature. Therefore, for higher Ge fractions, lower growth temperatures have to be applied. Moreover, for SiGe layers with high Ge content grown at high temperature, island growth and relaxation are more likely.

Murota (66) described the growth mechanism of low-temperature SiGe epitaxy based on a Langmuir-type adsorption scheme (67, 68). It is assumed that one SiH_4 or GeH_4 molecule is adsorbed at a single adsorption site, according to the Langmuir adsorption isotherm, and decomposes there. Then the SiH_4 and GeH_4 reaction rates R_{Si} and R_{Ge} are given by

$$R_{Si} = \frac{k_{Si}[k_1/(k_{Si} + k_{-1})]P_{SiH4}n_0}{1 + [k_1/(k_{Si} + k_{-1})]P_{SiH4} + [k_2/(k_{Ge} + k_{-2})]P_{GeH4}} \qquad 6.3$$

$$R_{Ge} = \frac{k_{Ge}[k_2/(k_{Ge} + k_{-2})]P_{GeH4}n_0}{1 + [k_1/(k_{Si} + k_{-1})]P_{SiH4} + [k_2/(k_{Ge} + k_{-2})]P_{GeH4}} \qquad 6.4$$

where k_{Si} and k_{Ge} are the SiH_4 and GeH_4 reaction rate constants, k_1 and k_2 the SiH_4 and GeH_4 adsorption rate constants, and k_{-1} and k_{-2} the SiH_4 and GeH_4 desorption rate constants, respectively; n_0 is the total adsorption site density; and P_{SiH4} and P_{GeH4} are the SiH_4 and GeH_4 partial pressures, respectively. The expressions for $K_{Si} = k_1/(k_{Si} + k_{-1})$ and $K_{Ge} = k_2/(k_{Ge} + k_{-2})$ denote the SiH_4 and GeH_4 effective adsorption equilibrium constants. Under some reasonable assumptions such as $[k_1/(k_{Si} + k_{-1})]P_{SiH4} >> [k_2/(k_{Ge} + k_{-2})]P_{GeH4}$, $k_{Si} \geqslant k_{-1}$ and $k_{Ge} \geqslant k_{-2}$, Eqs 6.3 and 6.4 become

$$R_{Si} = \frac{k_1 P_{SiH4} n_0}{1 + (k_1/k_{Si})P_{SiH4}} \qquad 6.5$$

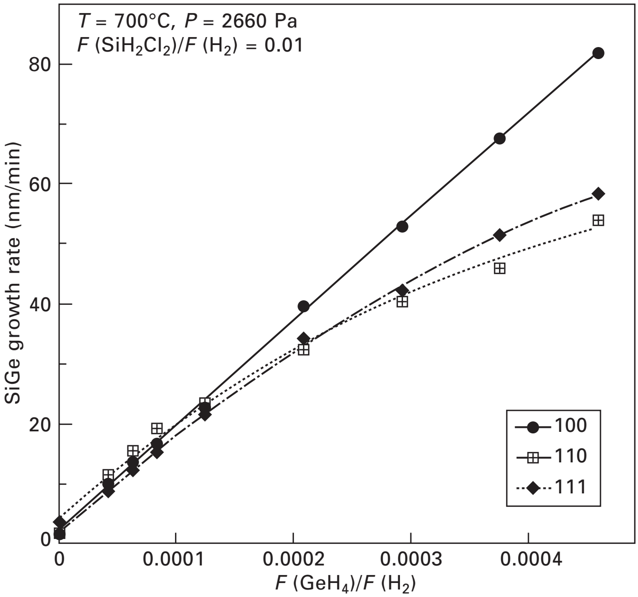
$$R_{Ge} = \frac{k_2 P_{GeH4} n_0}{1 + (k_1/k_{Si})P_{SiH4}} \qquad 6.6$$

The Ge fraction x is given by

$$x = \frac{R_{Ge}}{R_{Si} + R_{Ge}} = \frac{k_2}{k_1} \cdot \frac{P_{GeH4}}{P_{SiH4} + (k_2/k_1)P_{GeH4}} \qquad 6.7$$

Based on this reaction scheme, the experimentally found influence of the SiH_4 and GeH_4 partial pressures on the SiH_4 and GeH_4 reaction rate could be very well described.

The results for the growth of Si and SiGe reported so far have been obtained using Si(100) as the substrate. There is considerable interest in using other substrate orientations, especially for mobility enhancement for advanced MOSFETs. Hartmann *et al.* (69) investigated the growth kinetics of Si and SiGe on Si(100), Si(110) and Si(111) substrates using dichlorsilane/germane chemistry. In Fig. 6.8 the SiGe growth rates as a function of GeH_4/H_2 mass flow ratio for these three substrate orientations are shown as an



6.8 SiGe growth rate at 700°C on Si(100), Si(110) and Si(111) as a function of the GeH_4/H_2 mass flow ratio (69).

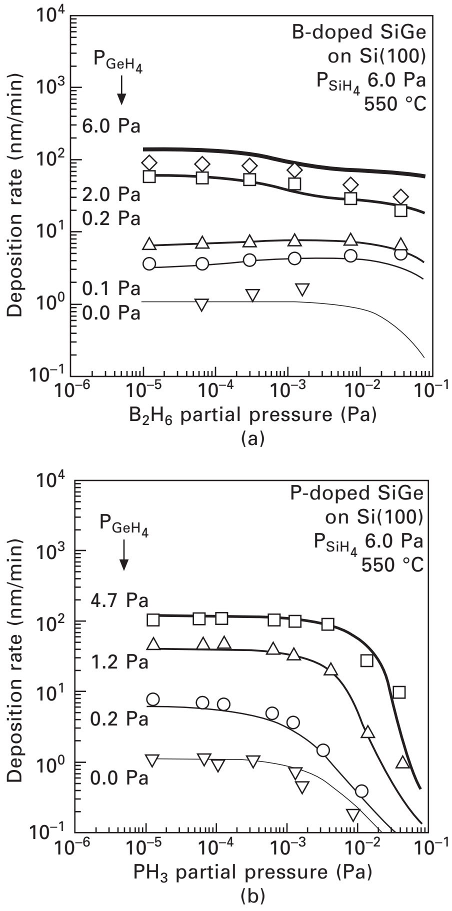
example. There is no significant difference in the low mass flow region (low Ge fraction in the film), whereas the growth rates on Si(110) and Si(111) are significantly lower in the high mass flow region (higher Ge fraction). The Ge fraction decreases when going from Si(100) to Si(111) to Si(110) substrates using the same process conditions.

6.4 Doping of silicon–germanium (SiGe)

6.4.1 *In-situ* doping of SiGe

Here B doping of SiGe by B_2H_6 and P doping by PH_3 are discussed (66). Figure 6.9 shows the deposition rate as a function of (a) the B_2H_6 and (b) the PH_3 partial pressures. A reduction of the deposition rate at higher GeH_4 partial pressures was investigated for B doping (Fig. 6.9(a)). There is no significant impact of the B_2H_6 addition on the Ge fraction (Fig. 6.10(a)). The B concentration increases nearly proportionally with the B_2H_6 partial pressure up to the 10^{22} cm^{-3} range (Fig. 6.11(a)).

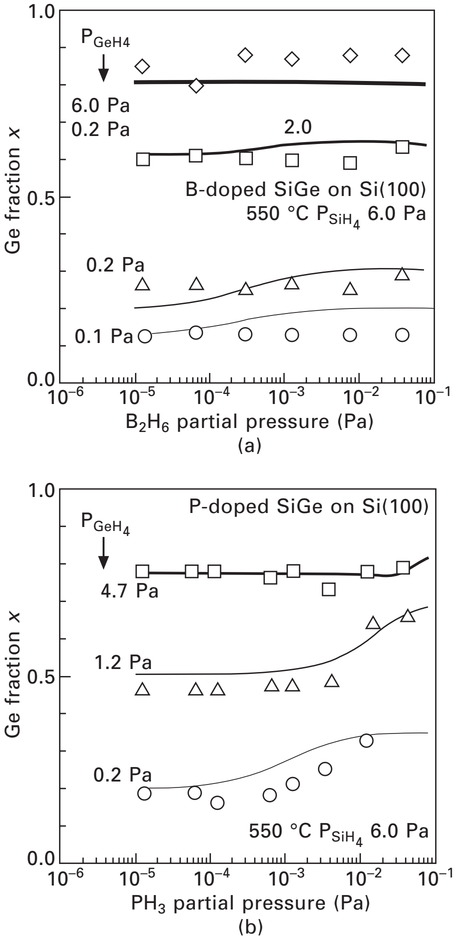
In the case of P doping, reduction of the deposition rate is investigated at the higher PH_3 partial pressures with increasing GeH_4 partial pressure (Fig. 6.9(b)). The reason for the reduced deposition rate is the lower density of free surface sites due to surface adsorption of PH_3, and also it is assumed that both the SiH_4 and GeH_4 molecules cannot react effectively on the P



6.9 Deposition rate as a function of (a) the B_2H_6 and (b) the PH_3 partial pressures (66).

or PH_x adsorbed sites. There is an increase in the Ge fraction for high PH_3 partial pressures (Fig. 6.10(b)). This effect is caused by larger reduction of the SiH_4 reaction as compared with the GeH_4 reaction on the surface. The P concentration increases up to a maximum value and then decreases with increasing PH_3 partial pressure (Fig. 6.11(b)). This effect can be explained assuming saturation of PH_3 adsorption and lower solubility of P on the Si–Ge and Ge–Ge pair sites than on the Si–Si pair site.

Figure 6.12 shows the relation between carrier concentration and dopant concentration (66). In the case of B-doped films, the carrier concentration is nearly equal to the B concentration up to about 2×10^{20} cm^{-3}, and it tends to saturate to about 5×10^{20} cm^{-3} at a B concentration below 10^{22} cm^{-3} (Fig. 6.12(a)). The same behavior was found to be independent of the Ge content

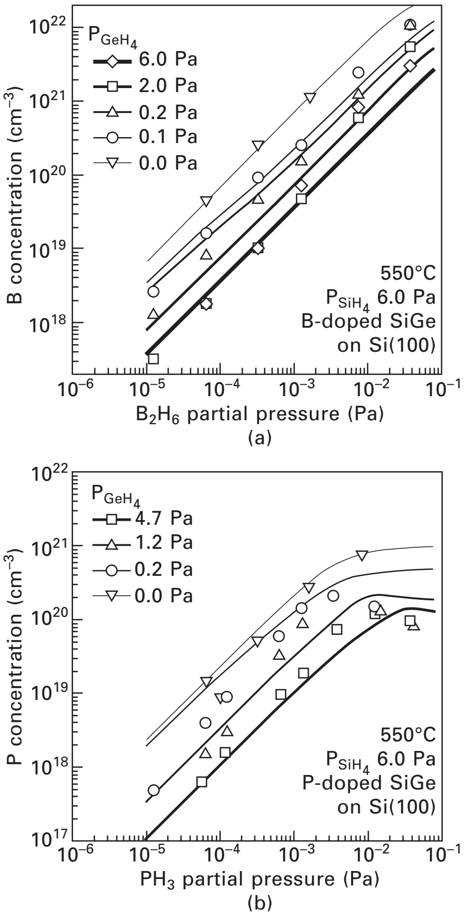


6.10 Ge fraction as a function of (a) the B_2H_6 and (b) the PH_3 partial pressures (66).

in the film. In the case of P-doped films for a Ge fraction higher than 0.5, the carrier concentration tended to saturate to about 10^{19} cm^{-3} at a higher P concentration up to about 2×10^{20} cm^{-3} (Fig. 6.12(b)), indicating electrically inactive P.

6.4.2 Heterojunction bipolar transistor (HBT) base doping and dopant profile control

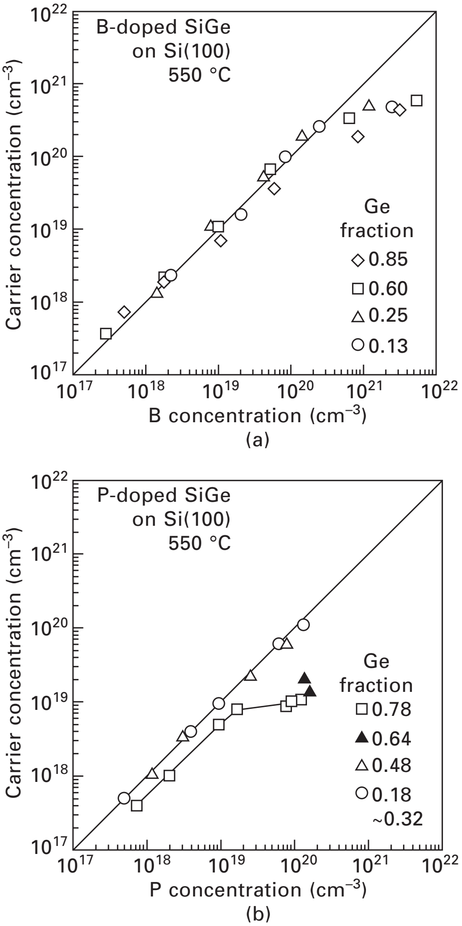
Doping of SiGe has been intensively investigated, especially in connection with HBT technology development. Complementary BiCMOS (CBiCMOS) SiGe technologies have been reported (70–72), demonstrating high levels



6.11 Dopant concentration in the $Si_{1-x}Ge_x$ films as a function of (a) the B_2H_6 and (b) the PH_3 partial pressures (66).

of npn and pnp RF performance. Within the CBiCMOS process technology a key enabler for high performance npn and pnp SiGe HBTs is the control of the vertical doping profiles, especially the base doping. This includes incorporation of high doping concentrations, the creation of steep doping profiles, and maintaining steep doping profiles during processing by controlling dopant diffusion and segregation. In this respect, results of B and P doping of SiGe for npn and pnp HBTs, respectively, are reviewed in this section. Moreover the impact of C incorporation on doping profile control is shown.

Considering the base doping of HBTs, key requests for optimized bipolar performance are the incorporation of a high concentration of electrically active dopant, the creation of steep doping profiles and the prevention of dopant



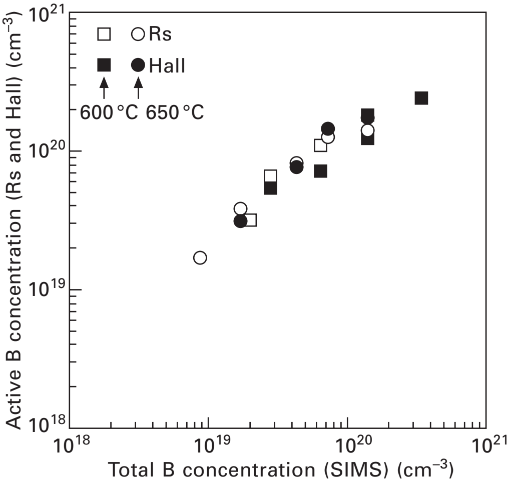
(a)

(b)

6.12 Carrier concentration in the $Si_{1-x}Ge_x$ films as a function of (a) the B and (b) the P concentrations (66).

profile broadening. Here B doping of SiGe npn HBT and P doping for pnp HBTs are reviewed (73). The doping of SiGe by C is shown to effectively suppress the diffusion of B and P in SiGe (74).

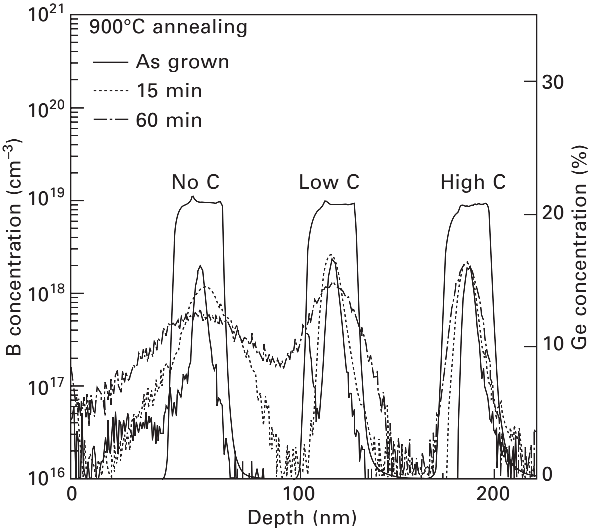
In Fig. 6.13 the total B concentration measured by SIMS incorporated into SiGe is compared with the electrically active concentration obtained from sheet resistance and Hall measurements. There is no significant difference between the two deposition temperatures used. The Ge fraction in the film was 20%. About the same concentration (within the error bars of the measurements) was obtained from electrical and SIMS measurements, indicating that the B is incorporated fully electrically active. As already shown in Fig. 6.11, a very high B concentration can be incorporated in the SiGe. For npn HBT a high doping concentration and prevention of dopant diffusion during



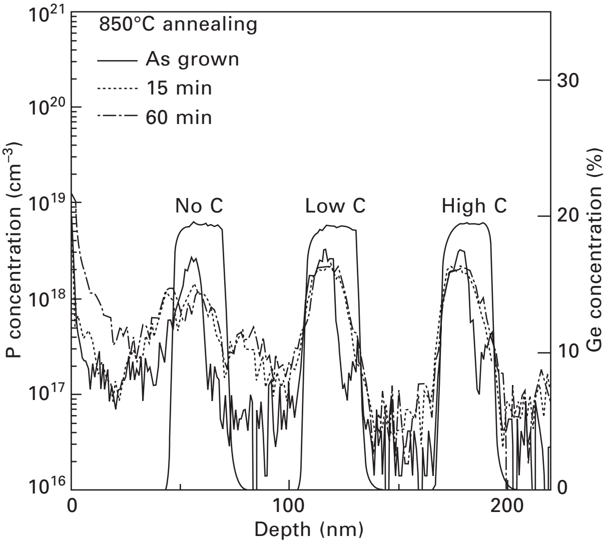
6.13 Comparison of B concentration in SiGe (20% Ge) measured by SIMS with electrically active B concentration obtained from sheet resistance and Hall measurements (73).

processing are essential. The incorporation of C into SiGe (in the order of 0.5%) can reduce the diffusion of B very effectively (23). A reduction of the B diffusivity by a factor of 20 has been reported for Si:C layers with C concentration of about 1×10^{20} cm^{-3} (75). Diffusion of P is suppressed in the same way, while diffusion of As and Sb is enhanced in C doped layers (76). Figure 6.14 shows the SIMS measurement of three B profiles incorporated in SiGe with 20% Ge as grown and after annealing. There was no C in the surface–near SiGe profile. For the other two SiGe profiles, low (about 3×10^{19} cm^{-3}) and high (about 1×10^{20} cm^{-3}) C concentrations were used. Very steep B profiles were obtained in as-grown SiGe and SiGe:C. The diffusion in the annealed samples depends on the C content and on the annealing conditions. In contrast to the profile without C, both the profiles with high and low C concentrations show limited broadening of the profile after 15 minutes of annealing at 900°C. At 900°C for 60 min annealing there is a small change of the B profile for high C concentration, whereas the two other profiles show increased broadening. By using a high C concentration there is minimum diffusion only for the B profile even after annealing at 900°C for one hour.

For the prevention of P base profile broadening of pnp HBT, C is used in the same manner. Figure 6.15 shows the result for P doping from the same experiment as was discussed above for B base doping. SIMS measurements of three P profiles incorporated in SiGe with 20% Ge as grown and after annealing are shown. There was no C in the surface–near SiGe profile. For the other two SiGe profiles, low (about 3×10^{19} cm^{-3}) and high (about 1



6.14 SIMS profile of B in SiGe and SiGe:C (20% Ge) for two C concentrations measured after deposition and after 900°C annealing for 15 min and 60 min (73).



6.15 SIMS profile of P in SiGe and SiGe:C (20% Ge) for two C concentrations measured after deposition and after 850°C annealing for 15 min and 60 min (73).

$\times\ 10^{20}$ cm^{-3}) C concentrations were used. There is a pile-up of P at the substrate site of the SiGe:C profiles and a less pronounced pile-up for the SiGe profile caused by autodoping from the previous P doping. For the high as well as for the low C concentration, there is limited broadening of the profiles after annealing at 850°C for 15 min and for 60 min. For the SiGe profile without C, the P profile is smeared out and there is an increased background doping level in the Si cap and spacer layers. Diffusion of P is suppressed by C doping in similar way as for B. Both B and P diffuse via an interstitial mechanism. The effect of C can be explained by a suppressed density of self-interstitials in C-doped SiGe.

Compared with the as-grown P profiles, the B profiles are much steeper. The broadening of the P profiles is caused by segregation during deposition of the SiGe stack. By reducing the temperature for the deposition of the undoped layer deposited on top of the P-doped SiGe, segregation can be effectively suppressed, resulting in steeper P profiles (77). The best value was achieved for the SiGe:C cap layer deposition at 500°C. In this case, a steepness of 6.3 nm/dec was obtained.

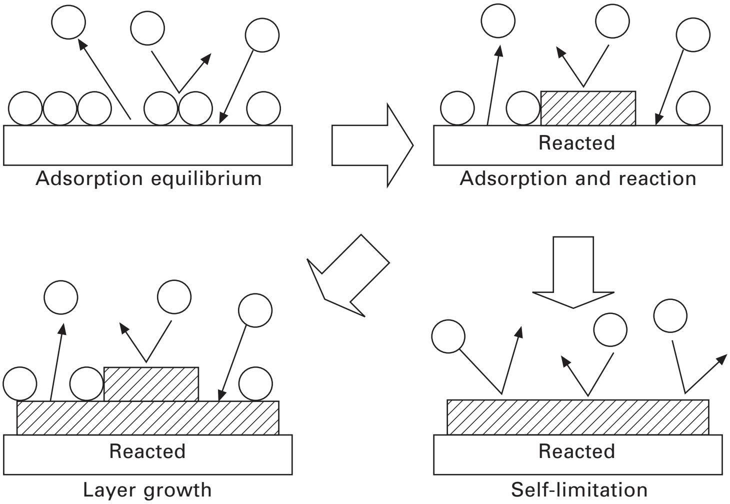
6.4.3 Atomic layer doping of SiGe (B, P)

In the following part the atomic layer processing approach (78–80) is discussed for the control of dopant incorporation on an atomic level and for the creation of very steep dopant profiles.

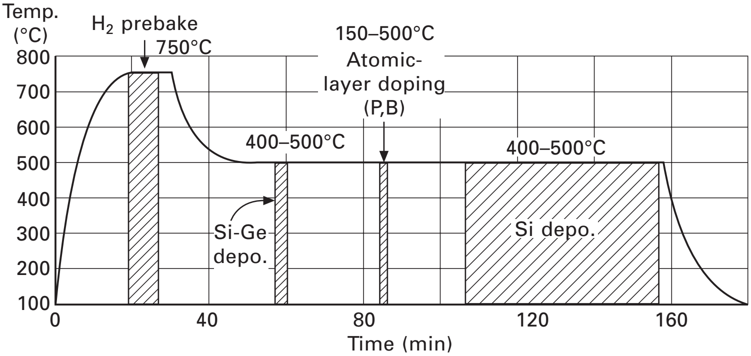
The concept of atomic-level processing for doping of Si, SiGe and Ge is based on atomic-order surface reaction control. By processing at low temperature the adsorption of reactant gases at the surface can be separated from the reaction process. In this way the process is controlled by the surface adsorption equilibrium only. The principle of surface reaction-limited processing is shown in Fig. 6.16. The surface of the wafer is exposed to the source gas molecules at low temperature. In the case of doping of SiGe the surface is exposed to dopant gas, PH_3 for P doping and B_2H_6 for B doping. Part desorption is taking place according to the adsorption–desorption equilibrium. The dopant molecules can react at the surface, resulting in dopant-covered surface areas. If the adsorption of dopant molecules occurs mainly at Si and Ge surface sites, it is limited by the availability of surface sites. If all surface sites are occupied, further adsorption is blocked and the process is self-limited. A maximum of one monolayer could be deposited in this case. Self-limited processing at low temperature is very beneficial for reliable and robust process technology.

If there is adsorption of molecules on sites which are occupied already, there is no self-limitation and several monolayers could be deposited at the surface.

The atomic layer processing approach has been demonstrated for the adsorption of different hydride gases on Si(100) and Ge(100) (81–90). Very low process temperatures (even room temperature) have been shown for epitaxy (Si, Ge) and doping (78–80, 91–93). The surface equilibrium


Adsorption equilibrium
Reacted
Adsorption and reaction
Reacted
Layer growth
Reacted
Self-limitation

6.16 Surface reaction-limited processing.



6.17 Process sequence for atomic layer doping using a batch reactor.

is determined by the exposure temperature and the partial pressures of the reactant gases. Moreover, the nature of the surface itself significantly affects the adsorption process. For example, the number of surface sites available for the adsorption process is different for a hydrogen-terminated surface compared to a surface which has been prepared to be free of hydrogen. This effect gives an additional parameter to control the adsorption of reactant gases at the surface.

Atomic layer doping can be performed both in single wafer and in batch reactors. In Fig. 6.17 the process sequence is illustrated for a batch reactor. Because very low temperatures are used for processing, very low moisture and oxygen levels in the reactor are required. The incorporation of dopants is performed without layer growth. Epitaxy is interrupted by switching off the source gases applied for Si, or SiGe layer growth. Then, the surface is exposed to PH_3/H_2 or B_2H_6/H_2, respectively. After surface exposure, the layer growth parameters are readjusted, and the epitaxial process is continued.

To increase the dopant coverage at low temperature, hydrogen termination of the surface can be prevented by switching from H_2 to N_2 at epitaxy temperature, cooling down in N_2 and performing the B_2H_6 and PH_3 exposure in N_2.

B atomic layer doping has been investigated for Si, SiGe and pure Ge with the aim of incorporating high dopant doses at low temperature, to create steep doping profiles and to control the B dose and location at an atomic level (78, 80, 92–94).

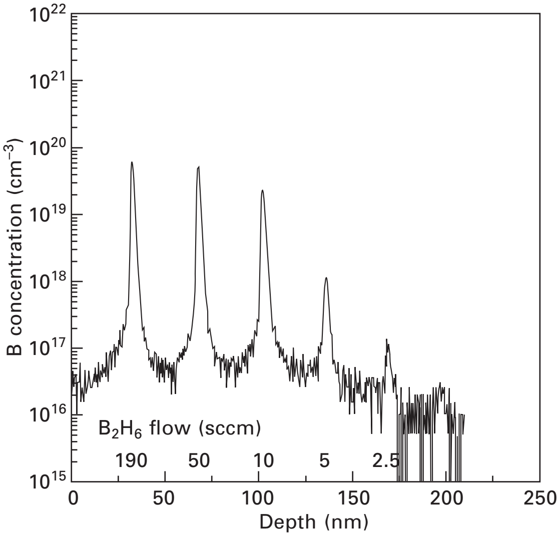
The B dose incorporated can be controlled by process parameters determining the surface adsorption equilibrium (partial pressure, exposure temperature). By increasing the partial pressure of B_2H_6 during the exposure step at the same temperature, the incorporated B dose is increased. It has been shown that the doping process is not self-limited at temperatures of 350°C and higher (80). Several monolayers of B have been incorporated (92, 93).

The high B concentration, incorporated during the exposure step, can only be explained by considering B adsorption at surface sites already occupied by B. This so-called B–B adsorption means clustering resulting in not electrically active B (95).

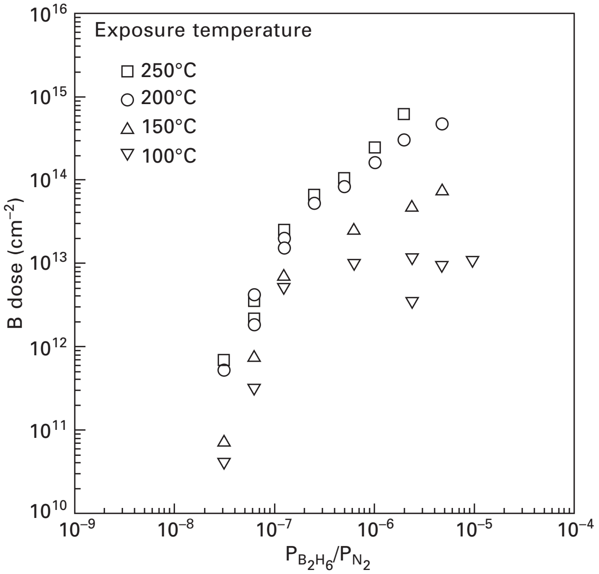
Tillack *et al.* (78, 93) investigated whether the nature of the B atomic layer doping process can be changed from non-self-limited to self-limited by decreasing the temperature. To illustrate this effect, in Fig. 6.18 the SIMS depth profiles of B peaks in SiGe at different B_2H_6 flows (partial pressures) at 100°C are shown. B_2H_6 exposure is performed on a hydrogen-free SiGe surface to increase the number of free surface sites for the adsorption of B. The B concentration increases with increasing B_2H_6 partial pressure and seems to be saturated at higher partial pressures. A very steep B profile (<1 nm/dec), in the order of the resolution limit of SIMS, is obtained.

The transition from self-limitation to non-self-limited behavior by changing the exposure temperature from 100°C to 250°C is shown in Fig. 6.19. For 200°C and 250°C exposure temperatures the B dose increases with increasing B_2H_6 partial pressure, reaching several monolayers of B. Clearly at these conditions there is no self-limitation of the process. By reducing the temperature to 100°C the character of the process is changed from non-self-limited to self-limited. The same trend was observed for B atomic layer doping of Ge.

For both SiGe and Ge this behavior could be explained considering a temperature-dependent adsorption on Si and Ge sites and B–B adsorption (78). At very low temperature, adsorption at Si and Ge surface sites seems to be favorable and B–B adsorption is suppressed. The adsorption is limited by



6.18 SIMS profile of B peaks in SiGe fabricated at different B_2H_6 partial pressures during exposure at 100°C in N_2 (78). B_2H_6 diluted to 250 ppm in H_2 is used.



6.19 B dose in SiGe as a function of the B_2H_6 to N_2 partial pressures for exposure between 100 and 250°C (78).

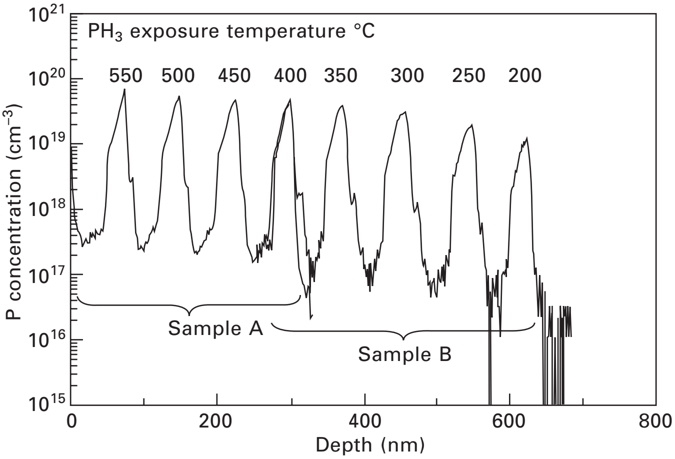
the availability of free Si or Ge surface sites. By increasing the temperature B–B adsorption seems to be dominating.

For P atomic layer doping, self-limitation of the process has been observed, allowing very precise dopant dose and location control. There have been indications that the partial pressure of H_2 during the adsorption step also has an influence on the P coverage (95). For high partial pressures of H_2, the adsorption of hydrogen could make a significant contribution, even if the equilibrium constant of the surface adsorption of H_2 is very small. Therefore, to control the P atomic layer doping process one has to consider the impact of the PH_3 and H_2 partial pressures during the exposure of the SiGe surface.

Figure 6.20 shows P profiles measured by SIMS for different exposure temperatures and exposure in N_2. High P concentrations have been obtained even at low temperature (200°C). But, compared with the very steep B profiles, the P peaks are much broader, mainly as a result of segregation. By lowering the temperature for processing on top of the P-doped layer, P segregation can be suppressed and steeper P profiles can be obtained.

6.5 Conclusion and future trends

Today low-temperature chemical vapor deposition of SiGe is an established technique that has been introduced in the semiconductor industry as well as in CMOS and in bipolar (HBT) technology. Despite the fact that the industrial success story of SiGe has been closely related to the usage of UHV batch deposition tools at the beginning, the dominating process tools are single wafer tools. There are numerous papers dealing with low-temperature epitaxy processing. Nevertheless, further investigations and development are required, especially regarding:



6.20 SIMS profile of P peaks in SiGe for PH_3 exposure temperature between 200°C and 550°C. Exposure on H-free surface by processing in N_2 (78).

- Very low temperature epitaxy for SiGe and Ge using alternative precursors
- Hardware improvements to secure very low moisture levels in the reactor to allow processing at very low temperature
- Alternative surface treatment to clean the wafer surface before epitaxy to lower the temperature for the hydrogen pre-bake or to prevent pre-bake
- Hardware improvements to improve the cost of ownership for low-temperature epitaxy.

Considering doping, high doping concentrations can be incorporated in SiGe, which are electrically active. Nevertheless, there is room for improvement, especially regarding:

- N type (P) doping of SiGe and especially Ge
- Creating and maintaining very steep and shallow doping profiles, e.g. for ultra-shallow junction formation
- Improved location, dose and uniformity control.

In this respect the atomic layer processing approach is very promising because it offers the opportunity to control processing at an atomic level and to obtain improved process robustness by using self-limited processing.

6.6 References

1. Cressler JD and Niu G, 2003, *Silicon–Germanium Heterojunction Bipolar Transistors*, Boston, MA: Artech House
2. Cressler JD, 2006, *Silicon Heterostructure Handbook*, Boca Raton, FL: CRC Press
3. Singh L, Harame DL and Oprysko MM, 2004, *Silicon Germanium Technology, Modeling, and Design*, Piscataway, NJ: IEEE Press
4. Pfeiffer UR, Öjefors E and Zhao Y, 2010, 'A SiGe quadrature transmitter and receiver chipset for emerging high-frequency applications at 160 GHz', *IEEE International Solid-State Circuits Conference*: 416; Öjefors E and Pfeiffer UR, 2010, 'A 650 GHz SiGe receiver front-end for terahertz imaging arrays', *IEEE International Solid-State Circuits Conference*: 430
5. Ghani T, Armstrong M, Auth C, Bost M, Charvat P, Glass G, Hoffmann T, Johnson K, Kenyon C, Klaus J, McIntyre B, Mistry K, Murthy A, Sandford J, Siberstein M, Sivakumar S, Smith P, Zawadzki K, Thompson S and Bohr M, 2003, 'A 90nm high volume manufacturing logic technology featuring novel 45nm gate length strained silicon CMOS transistors', *Technical Digest of the IEEE International Electron Devices Meeting*, Washington, DC: 978-980
6. Verdonckt-Vanderbroek S, Crabbe F, Meyerson BS, Harame DL, Restle PJ, Stork JMC and Johnson JB, 1994, 'SiGe channel heterojunction *p*-MOSFET's', *IEEE Trans. Elec. Dev.,* 41: 90–101
7. Miller KJ and Grieco MJ, 1962, 'Epitaxial silicon-germanium alloy films on silicon substrates', *J. Electrochem. Soc.,* 109: 70–71

8. Kasper E and Herzog HJ, 1997, ‘Elastic strain and misfit dislocation density in $Si_{0.92}Ge_{0.08}$ films on silicon substrates’, *Thin Solid Films*, 44: 357–370
9. Meyerson BS, 1992, ‘UHV/CVD growth of Si and Si:Ge alloys: Chemistry, physics, and device applications’, *Proc. IEEE,* 80: 1592–1608
10. Ahlgren DC, Gilbert M, Greenberg D, Jeng SJ, Malikowski J, Nguyen ND, Schonenberg K, Stein K, Groves R, Walter K, Hueckel G, Colavito D, Freeman G, Sunderland D, Harame DL and Meyerson B, 1996, ‘Manufacturability demonstration of an integrated SiGe HBT technology for the analog and wireless marketplace’, *Technical Digest of the IEEE International Electron Devices Meeting*, San Francisco, CA 859–862
11. Tillack B, Bolze D, Fischer G, Kissinger G, Knoll D, Ritter G, Schley P and Wolansky D, 1998, ‘SiGe heteroepitaxy for high frequency circuits’, *MRS Symposium Proceedings*, 525, *Rapid Thermal and Integrated Processing VII*: 379–384
12. Ritter G, Tillack B and Knoll D, 1995, ‘Successful preparation of high frequency HBT by integrated RTCVD processes’, *MRS Symposium Proceedings*, 387, *Rapid Thermal and Integrated Processing IV*: 341–346
13. Dutartre D, Warren P, Berbezier I and Perret P, 1992, ‘Low temperature silicon and $Si_{1-x}Ge_x$ epitaxy by rapid thermal chemical vapour deposition using hydrides’, *Thin Solid Films*, 222: 52–56
14. Sturm JC, Schwartz PV, Prinz EJ and Manoharan H, 1991, ‘Growth of $Si_{1-x}Ge_x$ by rapid thermal chemical vapor deposition and application to heterojunction bipolar transistors’, *J. Vac. Sci. Technol.* B, 9: 2011–2016
15. King CA, Hoyt JL, Noble DB, Gronet CM, Gibbons JF, Scott MP, Laderman SS, Kamins TI and Turner J, 1989, ‘Epitaxial growth of $Si_{1-x}Ge_x$/Si heterostructures by limited reaction processing for minority carrier device applications’, *MRS Symposium Proceedings*, 146, *Rapid Thermal Annealing/Chemical Vapor Deposition and Integrated Processing Symposium*: 71–82
16. de Boer WB and Meyer DJ, 1991, ‘Low temperature chemical vapor deposition of epitaxial Si and SiGe layers at atmospheric pressure’, *Appl. Phys. Lett.*, 58: 1286–1288
17. Meyer DJ and Kamins TI, 1992, ‘The deposition of Si–Ge strained layers from GeH_4, SiH_2Cl_2, SiH_4 and Si_2H_6’, *Thin Solid Films* 222: 30–33
18. Meyer DJ and Kamins TI, 1991, ‘Kinetics of silicon–germanium deposition by atmospheric-pressure chemical vapour deposition’, *Appl. Phys. Lett.*, 59: 178–180
19. Grützmacher DA, Sedgwick TO, Northrop GA, Powell AR and Kesan VP, 1993, ‘Very narrow SiGe/Si quantum wells deposited by low-temperature atmospheric pressure chemical vapour deposition’, *J. Vac. Sci. Technol. B*, 11: 1083–1087
20. Grützmacher DA, Sedgwick TO, Powell A, Tejwani M, Iyer SS, Cotte J and Cardone F, 1993, ‘Ge segregation in Si/Ge heterostructures and its dependence on deposition technique and growth atmosphere’, *Appl. Phys. Lett.*, 63: 2531–2533
21. Kühne H, Morgenstern T, Zaumseil P, Krüger D, Bugiel E and Ritter G, 1992, ‘Chemical vapour deposition of epitaxial SiGe thin films from SiH_4–GeH_4–HCl–H_2 gas mixtures in an atmospheric pressure process’, *Thin Solid Films*, 222: 34–37
22. Heinemann B, Knoll D, Fischer G, Krüger D, Lippert G, Osten HJ, Rücker H, Röpke W, Schley P and Tillack B, 1997, ‘Control of steep boron profiles in Si/SiGe heterojunction bipolar transistors’, *ESSDERC: Proceedings of the 27th European Solid-State Device Research Conference*, Stuttgart: 544–547
23. Osten HJ, Lippert G, Knoll D, Barth R, Heinemann B, Rücker H and Schley P, 1997, ‘The effect of carbon incorporation on SiGe heterobipolar transistor performance

and process margin', *Technical Digest of the IEEE International Electron Devices Meeting*, Washington, DC: 803–806

24. Knoll D, Heinemann B, Osten HJ, Ehwald KE, Tillack B, Schley P, Barth R, Matthes M, Park KS, Kim Y and Winkler W, 1998, 'Si/SiGe:C heterojunction bipolar transistors in an epi-free well, single-polysilicon technology', *Technical Digest of the IEEE International Electron Devices Meeting*, San Francisco, CA: 703–706
25. Osten HJ, Knoll D, Heinemann B, Rücker H and Tillack B, 1999, 'Carbon doped SiGe heterojunction bipolar transistors for high frequency applications', *Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting*, Minneapolis, MN: 109–116
26. Rücker H, Heinemann B, Barth R, Bauer J, Bolze D, Drews J, Fursenko O, Grabolla T, Haak U, Höppner W, Knoll D, Kuck B, Mai A, Marschmeyer S, Morgenstern T, Richter HH, Schley P, Schmidt D, Schulz K, Tillack B, Weidner G, Winkler W, Wolansky D, Wulf HE and Yamamoto Y, 2007, '130 nm SiGe BiCMOS technology with 3.0 ps gate delay', *Technical Digest of the IEEE International Electron Devices Meeting*, Washington, DC: 651
27. Smith FW and Ghidini G, 1982, 'Reaction of oxygen with Si(111) and (100): Critical conditions for the growth of SiO_2', *J. Electrochem. Soc.*, 129: 1301
28. Ghidini G and Smith FW, 1984, 'Interaction of H_2O with Si(111) and (100)', *J. Electrochem. Soc.*, 131: 2924
29. Meyerson BS, 1986, 'Low-temperature silicon epitaxy by ultrahigh vacuum/chemical vapor deposition', *Appl. Phys. Lett.*, 48: 797
30. Meyerson BS, Ganin E, Smith DA and Nguyen TN, 1986, 'Low temperature silicon epitaxy by hot wall ultrahigh vacuum/low pressure chemical vapor deposition techniques: Surface optimiztion', *J. Electrochem. Soc.*, 133: 1232
31. Murota J, Nakamura N, Kato M, Mikoshiba N and Ohmi T, 1989, 'Low-temperature silicon selective deposition and epitaxy on silicon using the thermal decomposition of silane under ultraclean environment', *Appl. Phys. Lett.*, 54: 1007–1009
32. Sedgwick TO, Berkenblitz M and Kuan TJ, 1989, 'Low-temperature selective epitaxial growth of silicon at atmospheric pressure', *Appl. Phys. Lett.*, 54: 2689
33. Gibbons JF, Gronet CM and Williams KE, 1985, 'Limited reaction processing: Silicon epitaxy', *Appl. Phys. Lett.*, 47: 721
34. Dutartre D, Warren P, Berbezier I and Perret P, 1992, 'Low temperature silicon and $Si_{1-x}Ge_x$ epitaxy by rapid thermal chemical vapour deposition using hydrides', *Thin Solid Films*, 222: 52
35. Pares G, Regolini JR, Mercier J, Dutartre D and Besahel D, 1990, 'Kinetics aspects of epitaxial silicon growth using disilane in a rapid thermal processing system', *J. Appl. Phys.*, 68: 4885
36. Isella G, Chrastina D, Rössner B, Hackbarth T, Herzog H-J, König U and von Känel H, 2004, 'Low energy plasma-enhanced chemical vapor deposition for strained Si and Ge heterostructures and devices', *Solid-State Electronics*, 48: 1317–1323
37. Sakuraba M, Muto D, Mori M, Sugawara K and Murota J, 2008, 'Very low-temperature epitaxial growth of silicon and germanium using plasma-assisted CVD', *Thin Solid Films*, 517: 10–13
38. Dutartre D, Loubet N, Brossard F, Vandelle B, Chevalier P, Chantre A, Monfray S, Fenouillet-Beranger S, Pouydebasque A and Skotnicki T, 2007, 'Si/SiGe epitaxy: a ubiquitous process for advanced electronics', *Technical Digest of the IEEE International Electron Devices Meeting*, Washington' DC: 689–692
39. Kunii Y, Inokuchi Y, Wang J, Yamamoto K, Moriya A, Hashiba Y, Kurokawa H and Murota J, 2006, 'Development of high-throughput batch-type epitaxial reactor', *ECS Transactions*, 3–7: 841–847

40. Matthews JW and Blakeslee AE, 1976, 'Defects in epitaxial multilayers, III. Preparation of almost perfect layers', *J. Crystal Growth*, 32: 265–273
41. Zaumseil P, 1994, 'A fast X-ray method to determine Ge content and relaxation of partly relaxed $Si_{1-x}Ge_x$ layers on silicon substrates', *Phys. Stat. Wol. (a)*, 141: 155–161
42. People R and Bean JC, 1986, 'Erratum: Calculation of critical layer thickness versus lattice mismatch for Ge_xSi_{1-x}/Si strained-layer heterostructures', *Appl. Phys. Lett.*, 49: 229
43. Van der Merwe JH, 1963, 'Crystal interfaces. Part II. Finite overgrowth', *J. Appl. Phys.*, 34: 123–127
44. Fitzgerald EA, Lee ML, Yu B, Lee KE, Dohrmann CL, Isaacson D, Langdo TA and Antoniadis DA, 2005, 'Dislocation engineering in strained MOS materials', *Technical Digest of the IEEE International Electron Devices Meeting*, Washington, DC: 519–522
45. Wolansky D, Tillack B, Blum K, Bolze KD, Glowatzki KD, Köpke K, Krüger D, Kurps R, Ritter G and Schley P, 1998, 'Low temperature clean for Si/SiGe epitaxy for CMOS integration of HBTs', *Electrochem. Soc. Proc.*, 98–1: 812–821
46. Wolansky D, Fischer GG, Knoll D, Bolze D, Tillack B, Schley P and Yamamoto Y, 2004, 'Impact of defects on the leakage currents of Si/SiGe/Si heterojunction bipolar transistors', *Solid State Phenomena*, 95–96: 249–254
47. Kern W and Puotinen DA, 1970, 'Cleaning solution based on hydrogen peroxide for use in semiconductor technology', *RCA*, 31: 187
48. Meyerson BS, Himpsel FJ and Uram KJ, 1990, 'Bistable conditions for low-temperature silicon epitaxy', *Appl. Phys. Lett.*, 57: 1034
49. Takagi T, Nagai I, Ishitani A, Kuroda H and Nagasawa Y, 1988, 'The formation of hydrogen passivated silicon single-crystal surfaces using ultraviolet cleaning and HF cleaning', *J. Appl. Phys.*, 64: 3516
50. Burrows VA, Chabal YJ, Higashi GS, Raghavachari K and Christman SB, 1988, 'Infrared spectroscopy of Si(111) surfaces after HF treatment: Hydrogen termination and surface morphology', *Appl. Phys. Lett.*, 53: 998
51. Chang KH and Sturm JC, 2007, 'Chlorine etching for *in-situ* low-temperature silicon surface cleaning for epitaxy applications', *ECS Transactions*, 6–1: 401–407
52. Liehr M, Greenlief CM, Kasai SR and Offenberg M, 1990, 'Kinetics of silicon epitaxy using SiH_4 in a rapid thermal chemical vapor deposition reactor', *Appl. Phys. Lett.*, 56: 629
53. Gu S, Zheng Y, Zhang R, Wang R and Zhong P, 1994, 'Ge composition and temperature dependence of the deposition of SiGe layers', *J. Appl. Phys.*, 75: 5382.
54. Hartmann JM, Loup V, Rolland G, Hollinger P, Laugier F, Vannuffel C and Semeria MN, 2002, 'SiGe growth kinetics and doping in reduced pressure-chemical vapor deposition', *J. Crystal Growth*, 236: 10–20
55. Thomson EV, Christensen C, Andersen CR, Pedersen EV, Egginton PN, Hansen O and Petersen JW, 1997, 'Cold-walled UHV/CVD batch reactor for the growth of $Si_{1-x}Ge_x$ layers', *Thin Solid Films*, 294: 72
56. Jung TG, Tsai CY, Huang GW and Wang PJ, 1994, 'Low-temperature epitaxial growth of silicon and silicon–germanium alloy by ultrahigh-vacuum chemical vapor deposition', *Jpn. J. Appl. Phys.*, 33: 240
57. Bozzo S, Lazzari JL, Coudreau C, Ronda A, Arnaudd Avitaya F, Derrien J, Mesters S, Holländer B, Gergaud P and Thomas O, 2000, 'Chemical vapour deposition of silicon–germanium heterostructures', *J. Crystal Growth*, 216: 171

58. Dutartre D, Warren P, Sagnes I, Badoz PA, Perio A, Dupuis JC and Prudon G, 1993, 'Epitaxy and doping of Si and $Si_{1-x}Ge_x$ at low temperature by rapid thermal chemical vapour deposition', *J. Vac. Sci. Technol. B*, 11: 1134
59. Beers AM and Bloem J, 1982, 'Temperature dependence of the growth rate of silicon prepared through chemical vapour deposition from silane', *Appl. Phys. Lett.*, 41: 153
60. Garone PM, Sturm JC, Schwartz PV, Schwartz SA and Wilkens BJ, 1990, 'Silicon vapour phase epitaxial growth catalysis by the presence of germane', *Appl. Phys. Lett.*, 56: 1275
61. Campell SA, Liu WH and Leighton JD, 1993, 'Effects of doping on growth in the dichlorosilane/germane system', *J. Vac. Sci. Technol. B*, 11: 1129
62. Knutson KL, Carl RW, Liu WH and Campell SA, 1994, 'A kinetics and transport model of dichlorosilane chemical vapour deposition', *J. Crystal Growth*, 140: 191
63. Kamins TI and Meyer DJ, 1991, 'Kinetics of silicon–germanium deposition by atmospheric-pressure chemical vapour deposition', *Appl. Phys. Lett.*, 59: 178
64. Sinniah K, Sherman MG, Lewis LB, Weinberg WH, Yates JT and Janda KC, 1989, 'New mechanism for hydrogen desorption from covalent surfaces: The monohydride phase on Si(100)', *Phys. Rev. Lett.*, 62: 567
65. Tillack B and Zaumseil P, 2006, in *Strained SiGe and Si Epitaxy*, ed. JD Cressler, Boca Raton, FL: CRC Press: 2.2-33–2.3-44
66. Murota J, 2001, 'Epitaxial growth techniques: Low temperature epitaxy', *Semiconductor and Semimetals*, 72: 127–149
67. Murota J and Ono S, 1994, 'Low-temperature epitaxial growth of $Si/Si_{1-x}Ge_x/Si$ heterostructure by chemical vapor deposition', *Jpn. J. Appl. Phys.*, 33: 2290–2299
68. Murota J, Sakuraba M and Matsuura T, 1999, 'High quality $Si_{1-x}Ge_x$ epitaxial growth by CVD', in *Defects in Silicon/1999*, ed. T Abe, WM Bullis, S Kobayashi, W Lin and P Wagner, *Electrochemical Society Proceedings*, 99–1: 189–202
69. Hartmann JM, Burdin M, Rolland G and Billon T, 2006, 'Growth kinetics of Si and SiGe on Si(100), Si(110) and Si(111) surfaces', *J. Crystal Growth*, 294: 288–295
70. El-Kareh B, Balster S, Leitz W, Steinmann P, Yasuda H, Corsr M, Dawoodi K, Dirnecker C, Foglietti P, Haeusler A, Menz P, Ramin M, Scharnagl T, Schiekofer M, Schober M, Sultz U, Swanson L, Tatman D, Waitschull M, Weijtmans JW and Willis C, 2003, 'A 5V complementary-SiGe BiCMOS technology for high-speed precision analog circuits', *Proceedings of the 2003 Bipolar/BiCMOS Technology Meeting (BCTM)*: 211
71. Heinemann B, Barth R, Bolze D, Drews J, Formanek P, Fursenko O, Glante M, Glowatzki K, Gregor A, Haak U, Höppner W, Kurps R, Marschmeyer S, Orlowski S, Rücker H, Schley P, Schmidt D, Schloz R, Winkler W and Yamamoto Y, 2003, 'A complementary BiCMOS technology with high speed *npn* and *pnp* SiGe:C HBTs', *Technical Digest of the 2003 International Devices Meeting (IEDM):* 117
72. Heinemann B, Barth R, Bolze D, Drews J, Formanek P, Grabolla T, Haak U, Köpke K, Kuck B, Kurps R, Marschmeyer S, Richter HH, Rücker H, Schley P, Schmidt D, Winkler W, Wolansky D, Wulf HE and Yamamoto Y, 2004, 'A low-parasitic collector construction for high-speed SiGe:C HBTs', *Technical Digest of the 2004 International Devices Meeting (IEDM):* 251
73. Tillack B, Heinemann B, Knoll D, Rücker H and Yamamoto Y, 2008, 'Base doping and dopant profile control of SiGe *npn* and *pnp* HBTs', *Appl. Surf. Sci.*, 254: 6013–6016
74. Rücker H, Heinemann B, Kurps R and Yamamoto Y, 2006, 'Dopant diffusion in SiGeC alloys', *ECS Transactions*, *3(7):* 1065–1075

75. Rücker H, Heinemann B, Röpke W, Kurps R, Krüger D, Lippert G and Osten HJ, 1998, 'Suppressed diffusion of boron and carbon in carbon-rich silicon', *Appl. Phys. Lett.*, 73: 1682; Rücker H, Heinemann B, Röpke W, Kurps R, Krüger D, Lippert G and Osten HJ, 1999, 'Erratum: Suppressed diffusion of boron and carbon in carbon-rich silicon', *Appl. Phys. Lett.*, 75: 147
76. Rücker H, Heinemann B, Bolze D, Knoll D, Krüger D, Kurps R, Osten HJ, Schley P, Tillack B and Zaumseil P, 1999, 'Dopant diffusion in C-doped Si and SiGe: Physical model and experimental verification', *Technical Digest of the IEEE International Electron Devices Meeting:* 109
77. Yamamoto Y, Tillack B, Köpke K and Kurps R, 2006, 'P doping control during SiGe:C epitaxy', *Thin Solid Films*, 508: 288
78. Tillack B, Yamamoto Y, Bolze D, Heinemann B, Rücker H, Knoll D, Murota J and Mehr W, 2006, 'Atomic layer processing for doping of SiGe', *Thin Solid Films*, 508: 279
79. Murota J, Sakuraba M and Tillack B, 2006, 'Atomically controlled processing for Group IV semiconductors by chemical vapor deposition', *Jpn. J. Appl. Phys.*, 45: 6767
80. Tillack B, Heinemann B and Knoll D, 2000, 'Atomic layer doping of SiGe – Fundamentals and device applications', *Thin Solid Films*, 369: 89–194
81. Murota J, Sakuraba M and Ono S, 1993, 'Silicon atomic layer growth controlled by flash heating in chemical vapor deposition using SiH_4 gas', *Appl. Phys. Lett.*, 62: 2353–2355
82. Sakuraba M, Murota J, Watanabe T, Sawada Y and Ono S, 1994, 'Atomic-layer epitaxy control of Ge and Si in flash-heating CVD using GeH_4 and SiH_4 gases', *Appl. Surf. Sci.*, 82–83: 354–358
83. Murota J, Sakuraba M, Watanabe T, Matsuura T and Sawada Y, 1995, 'Atomic layer-by-layer epitaxy of silicon and germanium using flash heating in CVD', *J. Phys. IV France*, 5, C5: 1101–1108
84. Watanabe T, Sakuraba M, Matsuura T and Murota J, 1997, 'Atomic-layer surface reaction of SiH_4 on Ge(100)', *Jpn. J. Appl. Phys.*, 36: 4042–4045
85. Sakuraba M, Murota J, Mikoshiba N and Ono S, 1991, 'Atomic layer epitaxy of germanium on silicon using flash heating chemical vapor deposition', *J. Crystal Growth*, 115: 79–82
86. Watanabe T, Ichikawa A, Sakuraba M, Matsuura T and Murota J, 1998, 'Atomic-order thermal nitridation of silicon at low temperatures', *J. Electrochem. Soc.*, 145: 4252–4256
87. Watanabe T, Sakuraba M, Matsuura T and Murota J, 1999, 'Separation between surface adsorption and reaction of NH_3 on Si(100) by flash heating', *Jpn. J. Appl. Phys.*, 38: 515
88. Shimamune Y, Sakuraba M, Matsuura T and Murota J, 2000, 'Atomic-layer adsorption of P on Si(100) and Ge(100) by PH_3 using an ultraclean low-pressure chemical vapor deposition', *Appl. Surf. Sci.*, 162–163: 390–394
89. Izena A, Sakuraba M, Matsuura T and Murota J, 1998, 'Low-temperature reaction of CH_4 on Si(100)', *J. Crystal Growth*, 188: 131–136
90. Takatsuka T, Fujiu M, Sakuraba S, Matsuura M and Murota J, 2000, 'Surface reaction of CH_3SiH_3 on Ge(100) and Si(100)', *Appl. Surf. Sci.*, 162–163: 156–160
91. Murota J, Matsuura T and Sakuraba M, 2002, 'Atomically controlled processing for group IV semiconductors', *Surf. Interface Anal.*, 34: 423–431
92. Tillack B, Yamamoto Y, Knoll D, Heinemann B, Schley P, Senapati B and Krüger

D, 2004, ‘High performance SiGe:C HBTs using atomic layer base doping’, *Appl. Surf. Sci.*, 224: 55

93. Tillack B, Heinemann B, Knoll D, Rücker H and Yamamoto Y, 2008, ‘Base doping and dopant profile control of SiGe *npn* and *pnp* HBTs’, *Appl. Surf. Sci.*, 254: 6013
94. Tillack B, Yamamoto Y and Murota J, 2004, ‘Atomic level control of SiGe epitaxy and doping’, *Proc. SiGe: Materials, Processing, and Devices: Proceedings of the 1st International Symposium*, 2004–07: 803
95. Tillack B, Zaumseil P, Morgenstern G, Krüger D and Ritter G, 1995, ‘Strain compensation in $Si_{1-x}Ge_x$ by heavy boron doping’, *Appl. Phys. Lett.*, 67: 1143

7
Strain engineering of silicon–germanium (SiGe) virtual substrates

K. SAWANO, Tokyo City University, Japan

Abstract: Strained Si/Ge channel devices with highly enhanced mobilities are formed on SiGe strain-relaxed buffer layers, called SiGe virtual substrates, hence crystal qualities of the virtual substrates crucially limit device performance. Relaxed layers essentially involve dislocations, meaning that the SiGe virtual substrate is far from the perfect crystal like a Si substrate. However, fabrication technologies have been intensively developed for two decades and nearly perfect crystal, applicable to high-performance devices, has become available. This chapter presents historical overviews and characteristic features for several fabrication techniques, such as compositionally graded buffer, LT buffer, ion-implantation buffer, etc.

Key words: SiGe virtual substrate, relaxed buffer, strained Si/Ge, dislocation.

7.1 Introduction

Large lattice-mismatch between Si and Ge (~4.2%) offers a concept of strain engineering, enabling us to arbitrarily manipulate and precisely control the lattice strain in device channels based on Si/Ge heterostructures. This concept makes the Si/Ge system remarkably attractive and promising for high performance CMOS applications since modification of band structures and accompanying significant mobility enhancements can be achieved by introducing proper strain states into the heterostructures. Adjustment of the strain in a very wide range is made possible by the use of $Si_{1-x}Ge_x$ alloys, where the lattice constant is continuously controllable by changing Ge concentrations (x). To create the strain-engineered channels on the Si-based platform, fully strain-relaxed SiGe alloys have to be formed on the Si substrate as buffer layers, on which the strained Si/Ge channel layers can be overgrown. Since such relaxed SiGe buffer layers are able to serve as 'substrates' possessing bulk-SiGe properties, they are called virtual substrates.

Strained Si channels with enhanced mobilities can be formed on SiGe virtual substrates with relatively low Ge concentrations (x = 0.2–0.3). In contrast, virtual substrates with higher Ge contents are needed for strained SiGe with higher Ge contents and strained Ge channels to be formed with 2D planar growth, since such channel layers largely mismatched to Si are likely to form islands with 3D growth when grown directly on Si or SiGe with low Ge contents. Moreover, high Ge content SiGe (or Ge) virtual substrates open a possibility of III–V material integration on the Si platform toward realization of opt-electronic integrated circuits (OEIC). Owing to this very beneficial and wide applicability, SiGe virtual substrates with a wide range of Ge concentrations have been targeted to date.

Formation of relaxed layers, i.e. SiGe virtual substrates, necessarily involves generation of misfit dislocations at a heterointerface between the SiGe and the Si substrate, which is always accompanied by generation of threading dislocations extending to the SiGe surface as well as surface roughness. Many efforts, therefore, have been made to create as perfect SiGe relaxed crystal as possible, mainly concentrating on minimizing the densities of threading dislocations and the level of surface roughness, as they critically determine the practical device applicability of the SiGe virtual substrate.

Since a direct growth of a constant composition SiGe layer on a Si substrate leads to a huge density of threading dislocation and large roughness, a variety of fabrication methods have been proposed and developed. In this chapter, typical techniques, such as compositionally graded buffer, low temperature buffer, ion implantation buffer, etc., are introduced and their fabrication methods, characteristic crystal qualities, applications, etc., are described.

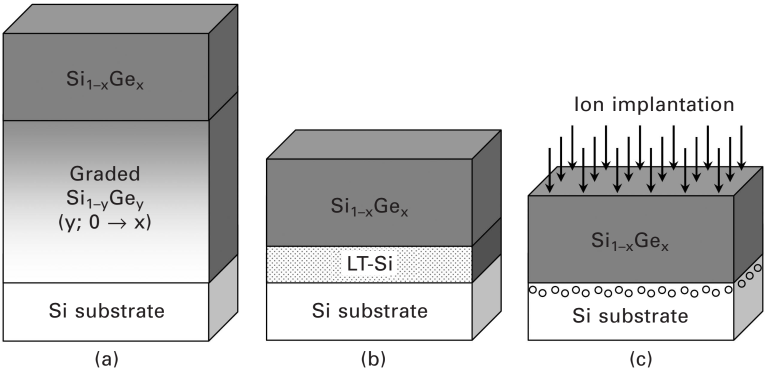
7.2 Compositionally graded buffer

7.2.1 Introduction

The strain accumulated in a SiGe film grown on a Si substrate can be relieved by introducing misfit dislocations at the heterointerface between the SiGe and the Si substrate. Hence, the larger strain requires the denser misfit dislocation arrays so as to be fully relaxed. The misfit dislocations gliding on one identical heterointerface interact with each other. Dislocation motion is highly impeded by the interaction, leaving threading segments extending to the surface. When the mismatch is relatively high, a large number of misfit dislocations must participate in the relaxation and as a result many interaction events take place, leading to large amounts of threading dislocations in the film, whereas films with lower mismatches can be grown with relatively small density of threading dislocations. Experimentally, the growth of SiGe film with lattice mismatches larger than ~1% has caused threading dislocation densities exceeding 10^8 cm^{-2}. One can expect from these facts that one solution is to increase the mismatch strain step by step in order to avoid introducing the largely mismatched heterointerface. This can be realized by inserting series of SiGe intermediate layers, where Ge concentrations are gradually increased with the thickness. This is called a compositionally graded buffer (Fig. 7.1(a)). A compositionally graded region is grown on a Si substrate, where the Ge concentration is gradually elevated linearly or step-wise. A typical grading rate is ~10%/µm, i.e. the Ge composition is increased by 10% with 1 µm thickness. This graded region is then capped by a uniform composition SiGe layer with the thickness of ~1 µm. Very promising and encouraging results had already been reported in 1991 by several groups [1–5] and very soon after, high mobility strained channels were demonstrated one after another [6–9] based on graded buffers with highly reduced threading dislocation densities. One can say that this graded method has dramatically opened the research field of strained Si/Ge channel heterodevices and still remains today the most standard, sophisticated and widely used technique among various methods. Comprehensive studies on the relaxation mechanisms concerning dislocation structures and surface morphologies for graded buffers were pursued mainly by two groups at IBM and AT&T [4, 5, 10–14].

7.2.2 Relaxation mechanism

Although the dislocation nucleation mechanism depends on growth conditions and consequently various mechanisms have been reported, it is commonly recognized that an essential mechanism, which can effectively reduce the threading dislocation density, is an enhanced threading dislocation gliding with very few trapping events. This can occur by distributing misfit dislocation arrays over the growth direction. Since the total number of misfit dislocations required for full strain relaxation is determined by the lattice mismatch, distributing them in the growth direction means reducing their density in



7.1 Schematics of various SiGe virtual substrates: (a) compositionally graded buffer, (b) low-temperature buffer, and (c) ion-implantation buffer.

a single lateral plane, which can decrease the probability of dislocation interaction occurring in an identical plane.

LeGoues *et al.* at IBM suggested that a modified Frank Read (MFR) source plays an important role in the relaxation mechanism for graded buffers as well as uniform composition SiGe buffer with small Ge contents [4, 11]. Possible origins of the MFR sources are pre-existing threading dislocations in Si substrates which move in orthogonal [110] directions on the SiGe/Si interface as misfit dislocations. Intersection of two dislocations having the same Burgers vector and gliding on two different (111) planes acts as the MFR source which produces dislocations repeatedly until the strain is sufficiently relaxed. It was proposed that threading segments can be effectively annihilated in this relaxation mechanism, leading to very low threading dislocation densities. A characteristic feature resulting from this mechanism is the formation of dislocation pile-ups deep within the Si substrate as a result of dislocation multiplication from the MFR sources. It was pointed out that pinning points of MFR sources should be sufficiently far apart for the MFR mechanism to effectively operate, and hence, random nucleation of high-density dislocations, which causes a large number of pinning points separated very closely from each other, has to be avoided. In this sense the MFR mechanism occurs only under very clean growth condition, where low energy sites for dislocation nucleation like particulates and/or defects are absent in the initial Si surface and growth process.

Surface roughness also produces detrimental dislocation nucleation sources. When a SiGe with relatively large mismatch grows pseudomorphically on a Si substrate, the accumulated strain is elastically and locally released via surface roughening like 3D island formation. Once the roughness forms, the stress is elastically concentrated in trench regions of the roughness and these regions become low energy sites for dislocation nucleation, bringing the relaxation. Such roughness-driven relaxation is, however, likely to result in high densities of threading dislocation due to large amounts of random nucleation sites. Hence surface roughening is regarded as a competing relaxation mechanism, which prevents MFR-related relaxation and consequently should be avoided.

It has been shown that the activation energy for strain-induced roughening decreases rapidly with an increase in misfit strain ε, varying as ε^{-4} [12]. In contrast, the activation energy for dislocation nucleation and multiplication varies much more slowly as ε^{-1}. This means that dislocation nucleation via surface roughening is favored at larger mismatch, whereas the MFR mechanism is favored at smaller mismatch. Therefore, the graded buffer layer consisting of low-mismatched layers relaxes via the MFR mechanism. In contrast, the uniform layer directly grown on Si without intermediate graded layers relaxes via roughening and subsequent random nucleation of dislocations. It can be said that dislocation nucleation in graded layers should begin while the SiGe with low Ge contents is growing. Unless the low Ge content SiGe layer relaxes sufficiently, the subsequent SiGe layers with higher Ge contents accumulate large strains like uniform layers without grading, and as a result roughening-driven relaxation accompanied by a high density of threading dislocations takes place. One can, therefore, say that an essential requirement for this method is sufficiently relaxing the strain step by step within the graded layers to keep the strain level low enough. In this sense, high growth temperatures and low grading rates are considered to be effective for reducing threading dislocations, which was suggested by Fitzgerald *et al.* at AT&T [5]. High growth temperatures can enhance the velocity of the dislocation motion and facilitate the strain relaxation. Low grading rates can suppress random nucleation of dislocations. Under these conditions, dislocations are able to glide long distances without few trapping events, leading to very long misfit dislocations and low threading dislocation density. Fitzgerald *et al.* employed a growth temperature of 900°C and a grading rate of 10% Ge/µm, and obtained very low threading dislocation densities, which were 4×10^5 cm^{-2} and 3×10^6 cm^{-2} for $Si_{0.77}Ge_{0.23}$ and $Si_{0.50}Ge_{0.50}$ graded buffers, respectively [5]. In contrast, LeGoues *et al.* reported a very low density, ~10^4 cm^{-2}, for the $Si_{0.75}Ge_{0.25}$ graded buffer with a relatively high grading rate (50% Ge/µm) and low growth temperatures (500°C) [4]. They explained that lowering growth temperature can suppress the surface roughening owing to reduced migration of adatoms and allow the relaxation by the MFR mechanism, where annihilation of threading dislocations plays a crucial role for realizing the low density of threading dislocation [13].

7.2.3 Crosshatch morphology

Many studies have focused on surface morphology as well as dislocation structures since these two features are strongly correlated with each other. From the viewpoint of application to strained channel devices, large surface roughness not only brings significant difficulty in the lithography process but also directly affects carrier transport in the channel formed on rough buffers.

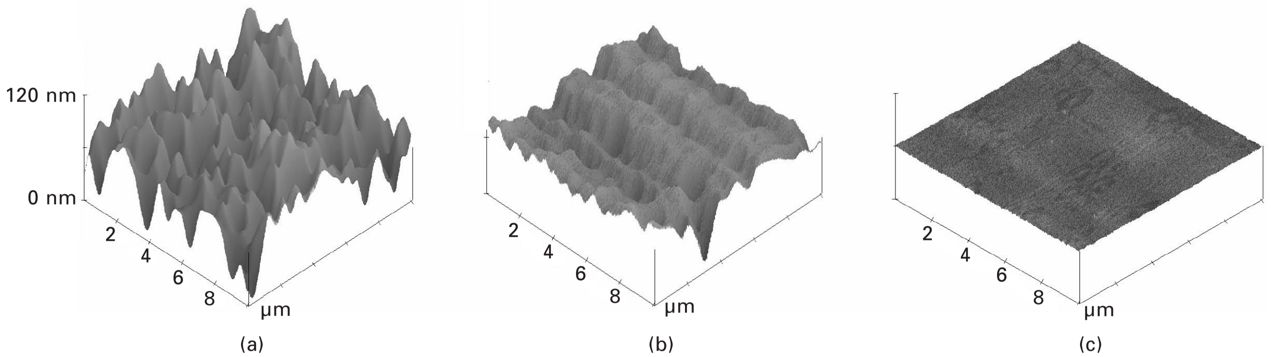
As described above, when a SiGe layer has a large mismatch strain, elastic relaxation causes large surface roughening. Even at a small mismatch strain, where such relaxation via roughening can be avoided, strain relaxation through the formation of misfit dislocations also leads to surface roughness. A characteristic feature of this roughness morphology is the appearance of a so-called crosshatch pattern consisting of periodic surface ridges aligned along orthogonal [110] directions. Since the alignment directions are the same as those of misfit dislocations formed in graded regions, it is clear that the misfit dislocations are responsible for the roughness formation.

Mechanisms of the crosshatch formation have been argued, and two main mechanisms are considered to operate. One is via surface steps created as a result of misfit dislocation generation [15, 16]. A generated misfit dislocation glides on a (111) plane forming a step on the surface at least the size of Burgers vector. As the relaxation proceeds, multiple misfit dislocations gliding on a common (111) plane form multiple surface steps, resulting in the periodic ridges along (110) directions. This mechanism dominates the crosshatch formation for slightly relaxed SiGe with the thickness comparable to or slightly higher than the critical thickness. For thicker SiGe, by contrast, another mechanism comes into play. The local strain field associated with the presence of misfit dislocations becomes a more dominant cause of the roughening. This mechanism was proposed by Fitzgerald and co-workers [7, 17] based on the fact that the surface roughness of SiGe buffers strongly depends on the grading rates even though Ge compositions as well as relaxation ratios of the buffers are equal, that is, the total number of misfit dislocations must be the same. It was also argued that the roughness heights are too large to be explained solely by multiple surface steps induced by misfit dislocations. For these reasons, they predicted that an inhomogeneous strain field distribution is formed at the SiGe surface due to nonuniform misfit dislocation arrays in the graded region. Once the inhomogeneous strain field arises at the surface of a growing SiGe layer, the surface roughening occurs to minimize the total free energy by trading strain energy for surface energy, resulting in the large crosshatch roughness. As the grading rate is increased, the misfit dislocation nucleation begins earlier due to lower critical thickness, and misfit dislocation arrays are located closer to the surface, making the stronger inhomogeneous strain field at the surface and leading to more severe roughening. The strain distribution as well as related roughness formation was further studied by means of Raman mapping and CMP, as will be shown below.

Severe crosshatch roughness brings an unfavorable increase in threading dislocation density [18, 19]. In a region of a relatively deep trench of the crosshatch surface, threading dislocation motion is strongly impeded and arrested since the surface depression together with strain fields due to orthogonal misfit dislocations reduces a driving force for the dislocation motion in this region. A number of threading dislocations get trapped and form dislocation pile-ups along the trench. Moreover, the growth rate is locally reduced around the trench where the dislocations terminate, making the trench deeper as growth proceeds. The deeper trench blocks more likely the dislocation motion and incorporates more threading dislocations, forming more pile-ups. Since the trapped dislocation is no longer able to contribute to the relaxation, additional dislocation nucleation is required, leading to larger threading dislocation densities. This mechanism can explain the fact that SiGe graded buffers with higher Ge contents contain higher densities of threading dislocations even though the grading rate is kept sufficiently low. Suppressing the density of the pile-ups is, therefore, quite essential for creating high-Ge-content SiGe virtual substrates with very low threading dislocation densities. It was suggested that the employment of off-cut Si substrates can aid reducing the pile-up density [18, 19] since blocking of the dislocation motion is effectively avoided owing to a nonparallel misfit dislocation network formed when SiGe graded buffers are grown on the off-cut substrates. A more effective way is directly removing the crosshatch trenches, i.e. the sources of pile-ups. For the purpose of eliminating such pile-ups, surface polishing via a chemical mechanical planarization (CMP) technique was attempted by Currie *et al.* [20]. They carried out surface polishing of the graded SiGe buffer when the Ge content was graded up to 50%, and continued the growth of the graded layer up to 100% Ge. The polishing eliminated the roughness, enabling the threading dislocations trapped in pile-ups to move and contribute to the relaxation. Consequently, additional dislocation nucleation was unnecessary and the density of threading dislocation was reduced during the regrowth. As a result, the density as low as 2.1×10^6 cm^{-2} was achieved for the pure Ge virtual substrate.

CMP of SiGe was further examined, focusing on the formation of an ultra-smooth surface for SiGe virtual substrates with various Ge concentrations [21, 22]. Adopting a similar CMP process as has been established for Si wafer production and the proper post-CMP cleaning method, SiGe virtual substrates with a wide range of Ge concentrations can be planarized down to a root-mean-square (rms) roughness less than 1 nm in spite of initial amplitudes of the roughness. Figure 7.2 shows the surface AFM images of the $Si_{0.7}Ge_{0.3}$ virtual substrates before and after CMP. Although the rms surface roughness is as large as 20 nm before CMP (Fig. 7.2(a)), as the polishing proceeds ridge regions of the roughness are selectively polished and the roughness rapidly decreases. Finally, an ultra-smooth surface is realized with rms roughness below 0.5 nm.

Post-CMP cleaning is a critically important issue for successful regrowth of device channel structures. Any contamination, especially polishing particles in slurry, has to be completely removed by appropriate cleaning procedures. The standard cleaning method established for the cleaning of the Si substrate is not always suitable for the SiGe cleaning as it is. In many cases, after post-CMP cleaning, the crosshatch roughness appeared again on the surface even though it had once been completely eliminated by the polishing. A laterally non-uniform etching rate of the SiGe surface is responsible for this roughening. This non-uniformity comes from the aforementioned inhomogeneous strain field distribution with crosshatch morphology. In the presence of such strain distribution, which is considered to be unavoidable for graded buffers, surface etching has to be suppressed as much as possible while maintaining efficient cleaning effects. For removal of particles from the surface, $NH_4OH + H_2O_2$ cleaning, known as APM



7.2 The surface AFM images of the $Si_{0.7}Ge_{0.3}$ virtual substrates (a) before, (b) after partial CMP, and (c) after total CMP.

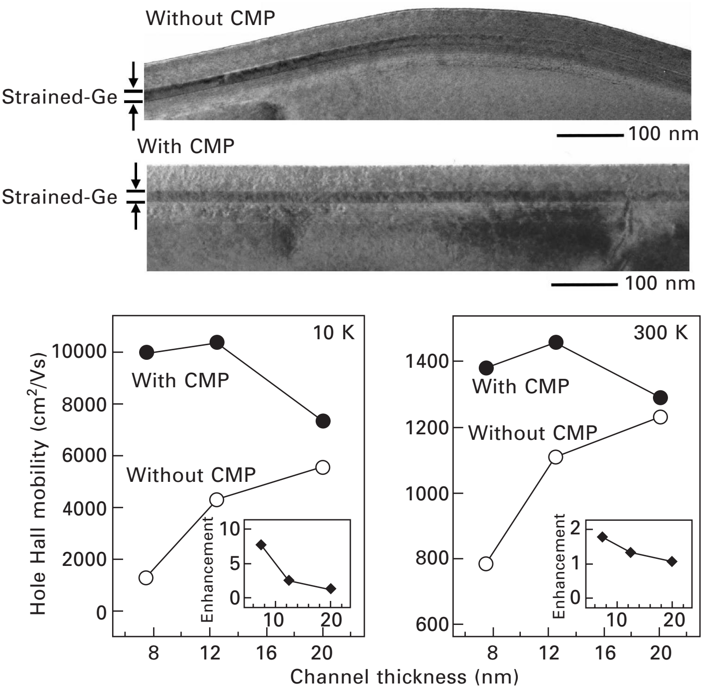
cleaning, is the most useful and effective method. One should take into account, here, that NH_4OH etches Si whereas H_2O_2 etches Ge, hence the degree of etching and accompanying surface roughening of SiGe depends on the ratio of NH_4OH to H_2O_2 as well as on Ge concentration. The ratio was optimized for various Ge concentrations to minimize roughening without reducing cleaning efficiency [22]. In addition, the cleaning temperature should be as low as possible because high-temperature cleaning enhances the etching rate. Room-temperature cleaning was found to work effectively with the help of ultrasonic treatments.

How the planarization of SiGe virtual substrates affects transport properties of the device formed on them, especially carrier mobility, is dependent on structures and measurement temperatures. For application to strained Si channels, significant improvements in low-temperature mobility were reported while effects on room-temperature mobility were not remarkable [23–25]. This is because the phonon scattering is dominant at room temperature, making the scattering caused by the roughness negligible. In contrast, for applications to strained Ge channels [26], the planarization effect is more significant since SiGe virtual substrates with higher Ge concentrations, which are generally used for strained Ge channels, involve much larger roughness indispensably. Mobilities with and without CMP are compared for the strained Ge grown on $Si_{0.35}Ge_{0.65}$ buffer layers. Figure 7.3 shows hole Hall mobility at 10 K and 300 K as a function of the channel thickness. At both temperatures, it is clearly found that the mobility drastically decreases with decreasing channel thickness for the sample without CMP, while no reduction is seen and high mobility is maintained for the CMP sample. Successful elimination of tilting of the channel layer as seen in XTEM images and resultant improvement of the heterointerface quality may contribute to the mobility enhancement, particularly for thin channels. The mobility enhancement factor owing to CMP, as shown in insets of Fig. 7.3, reaches 8 and 1.8 at 10 K and 300 K, respectively, at the channel thickness of 7.5 nm.

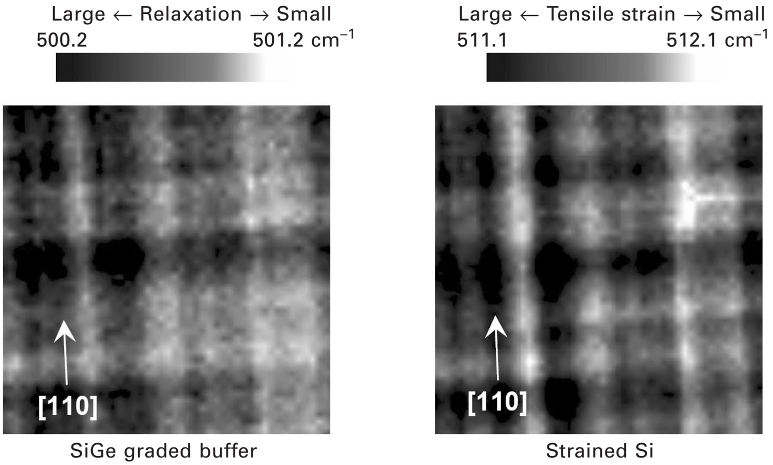
7.2.4 Strain field fluctuation

The problem still remaining even after the surface planarization is the presence of large numbers of misfit dislocations within graded layers. The underlying misfit dislocations intrinsically induce strain field spreading all over the SiGe layer and possibly degrade properties of overgrown devices. As discussed above, non-uniform strain distribution is considered to remain on the fully relaxed SiGe surface due to inhomogeneous misfit dislocation arrays, which was predicted by calculations [7].

Micro Raman spectroscopy was utilized to directly observe this strain distribution [27]. Figure 7.4 shows Raman mapping images of the graded $Si_{0.7}Ge_{0.3}$ buffer layer and the overgrown strained Si layer. Although this



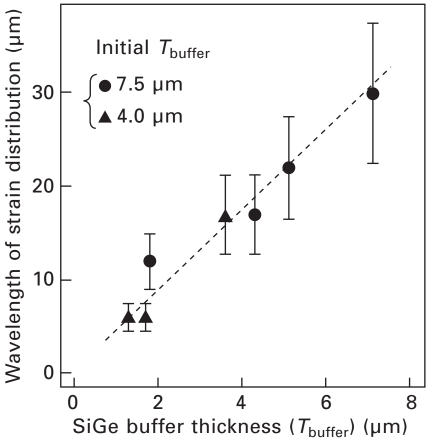
7.3 Cross-sectional TEM images of strained Ge channel layers grown on $Si_{0.35}Ge_{0.65}$ virtual substrates with and without CMP, and thickness dependence of hole Hall mobility for the channels with and without CMP at 10 and 300 K.


Large ← Relaxation → Small
500.2
501.2 cm^{-1}
Large ← Tensile strain → Small
511.1
512.1 cm^{-1}
[110]
[110]
SiGe graded buffer
Strained Si

7.4 Raman mapping images for the structure of strained-Si/$Si_{0.7}Ge_{0.3}$ virtual substrate.

SiGe layer was planarized by CMP and consequently has no crosshatch surface roughness, crosshatch-like contrasts, i.e. strain field distribution, along two orthogonal (110) directions are obviously seen, indicating that the inhomogeneous strain field is distributed in a similar way to the crosshatch patterns. Note that the SiGe was almost fully relaxed owing to its thickness being far beyond the equilibrium critical thickness and to high growth temperature. This indicates that the strain field induced by misfit dislocations extends to the SiGe surface and causes the strain fluctuation of the strained Si layer. Moreover, it is noticeable that the SiGe buffer and the strained Si layers exhibit a similar strain distribution. Since the strained Si is pseudomorphically grown on the SiGe buffer, the strain variation of the underlying SiGe directly causes that of the overgrown strained Si. Quantitatively, the Raman mapping results reveal that the tensile strain of the strained Si ranges from 0.0109 to 0.0125 (Fig. 7.4), meaning that approximately 13% of the strain fluctuates. This strain variation, which possibly causes fluctuation and/or deterioration of device properties, should be taken into account for device applications and preferably reduced. One simple solution to this problem is increasing SiGe thickness.

It was found that a lateral wavelength of the strain distribution monotonically increases with increase in SiGe thickness (Fig. 7.5) [28]. Surprisingly, the fluctuation was revealed to remain in SiGe with the thickness as large as 7 μm. In other words, such a large thickness is necessary to realize uniform strain distribution. Such a large thickness is, however, practically not favorable. Another way to reduce the strain variation besides increasing the thickness



7.5 Wavelength of strain distribution as a function of the thickness of SiGe virtual substrate. The thickness was changed by surface polishing, and initial T_{buffer} means the buffer thickness before polishing.

is altering dislocation structures by introducing defects as dislocation sources via low-temperature growth or ion bombardment, which will be described below.

An effect of the strain field distribution on the surface roughening was also investigated through combination with CMP [27, 28]. When a SiGe layer was homoepitaxially grown on a planarized SiGe buffer with the same Ge concentration, the crosshatch roughness was found to reappear on the surface. The morphology was completely the same as that observed with Raman mapping. This is a clear verification that the crosshatch roughness is formed by the inhomogeneous strain field distribution. The surface diffusion of adatoms is considered to be modulated by the local strain field.

7.3 Low-temperature buffer

Although the compositionally graded buffer method has been the most effective way to reduce the threading dislocation density, it has several drawbacks from the device application point of view. One is that the growth of very thick SiGe layers, more than several μm, is essential to efficiently reduce the density of threading dislocation. The large thickness is time and material consuming and could affect device properties. Due to the narrower bandgap and lower thermal conductivity of the SiGe alloy than those of Si, the larger off-leak currents and self-heating effects in MOS devices could be caused by the larger SiGe thickness. Secondly, the unavoidable large surface roughness requires an additional planarization process with CMP. Moreover, an inhomogeneous strain field resides due to non-uniform distribution of underlying misfit dislocations at the surface even after a thick layer is grown, and planarized. Therefore, modification of dislocation structures via control of the generation and multiplication of dislocations is essential. Intentional introduction of defects into the heterostructures is one solution, since they can act as dislocation nucleation sources [29]. Such defects can be induced by low-temperature crystal growth or ion bombardment. Both methods, known as low-temperature (LT) buffer and ion-implantation buffer, respectively, have been intensively studied since the mid-1990s following the success of the graded method. Ion implantation method is discussed in Section 7.4.

The LT buffer normally consists of a Si buffer layer grown on a Si substrate at low temperature (LT-Si) and a following SiGe layer grown at higher temperature (Fig. 7.1(b)). Since surface migration of Si adatoms is largely restricted due to the low growth temperature, point defects are induced in the LT-Si layer. During subsequent SiGe growth, dislocation generation is well accelerated owing to the point defects acting as low energy sites for dislocation nucleation. With optimal conditions, the LT-Si layer confines dislocations and prevents threading dislocations from extending to the surface. Important growth parameters to be optimized are the growth temperature and the thickness of the LT-Si layer. The temperature range has to be low enough to generate high-density point defects but high enough to retain crystal growth. Typical growth temperatures and the thicknesses of the LT-Si so as to obtain high-quality SiGe buffers are 350–450°C and 50–200 nm, respectively.

The LT buffer method was first applied to SiGe buffer fabrication by Chen *et al.* in 1996 [30]. They grew a 50 nm thick LT-Si layer at 400°C and subsequent 300 nm thick $Si_{0.76}Ge_{0.24}$ at 550°C, which resulted in a threading dislocation density of ~10^6 cm^{-2}. Surprisingly, this first attempt was quite successful and the employed growth parameters were almost optimal. A year later, they extended the Ge concentration up to 30% and achieved a threading dislocation density in the order of 10^5 cm^{-2} [31]. It was found that the relaxation ratio monotonically increased with the SiGe thickness and reached 90% at a thickness of 500 nm, which was much thinner than that of the comparable graded buffers. A remarkable feature was the very small surface roughness. The rms roughness obtained was only 1.2 nm, significantly lower than those of graded buffers. A little earlier, in 1996, Linder *et al.* also reported on the use of the LT buffer [32], where unusually gas-source MBE was used for the LT-Si buffer growth, and a $Si_{0.85}Ge_{0.15}$ with threading dislocation density lower than 10^4 cm^{-2} was obtained.

After these successful early works, growth conditions were reexamined to attain higher quality films and to elucidate the strain relaxation mechanism. The effects of LT-Si buffer thickness were systematically investigated by Lee *et al.* [33]. It was found that degrees of strain relaxation of 300 nm thick $Si_{0.7}Ge_{0.3}$ exceeded 90% independently of the LT-Si thickness, whereas the threading dislocation density was found to decrease markedly with increase in the thickness from 50 to 200 nm and to be equivalently low level (~10^6 cm^{-2}) from 200 to 300 nm. It was suggested that the thicker LT-Si buffer layer contains a larger amount of vacancies that effectively block the propagation of threading dislocations into the SiGe layer and confine dislocation half-loops in the LT-Si layer. Luo *et al.* studied the effects of the growth temperature as well as the thickness of the LT-Si buffer [34]. An optimal LT-Si growth temperature was shown to be 400°C. At higher temperatures, the numbers of induced point defects are not sufficient to facilitate the strain relaxation. On the other hand, at lower temperatures, although many point defects are generated, they aggregate and form stacking faults before SiGe growth, leading to large surface roughness. Luo *et al.* also showed that an optimal LT-Si thickness was 200 nm in terms of threading dislocation density and surface roughness, which agreed with the report by Lee *et al.* [33]. Eventually, 500 nm thick $Si_{0.7}Ge_{0.3}$ with a threading dislocation density as low as 1.5×10^5 cm^{-2} was obtained with the optimized conditions. Luo *et al.* also experimentally revealed that the LT-Si layer has so-called compliant effects, that is, the LT-Si layer becomes tensilely strained due to the stress provided by the overgrown compressively strained SiGe, and the lattice mismatch between the LT-Si and SiGe layers was reduced, contributing to reduction of threading dislocation density.

States of defects induced in the LT grown layer were investigated with the use of positron annihilation spectroscopy [35, 36]. It was shown that, in the case of low temperature (200–400°C) epitaxy of Si and SiGe, the layer contained defects of the vacancy type with a concentration of 10^{17}–10^{18} cm^{-3}. At temperatures of 350–400°C, which can bring high quality SiGe buffers, they were found to form clusters of larger sizes. These vacancy clusters are reasonably considered to contribute to reduction of threading dislocation densities.

To obtain SiGe buffers with higher Ge concentrations with the LT buffer method, some modifications are necessary. Actually, straightforward application of the manner employed for low Ge contents (Ge < 30%) to high Ge contents (Ge > 30%), that is, growth of a single SiGe layer with a constant Ge content on the LT-Si buffer, failed to effectively reduce threading dislocation density. Generation of dislocation half-loops from the surface was shown to be responsible for the large threading dislocation density [37]. To overcome this problem, step-wise elevation of the Ge composition was attempted, where LT grown layers were inserted at each step [36–38]. For example, in the case of a two-step LT buffer, a 500 nm thick SiGe with ~30% Ge is first grown on the LT-Si buffer. Next, LT growth of SiGe with the same Ge contents as first grown SiGe, i.e. 30%, is performed and followed by growth of 500 nm thick SiGe with ~60% Ge. This concept is analogous to the graded buffer method, but since Ge compositions jump by about 30% at each step with only 500 nm thickness, the total thickness is markedly lower than that of graded buffers. With the three-step LT method, Peng *et al.* fabricated the SiGe buffer with the Ge content as high as 90% and achieved threading dislocation density lower than 5×10^6 cm^{-2} with a total thickness below 1.7 μm [38]. A detailed discussion on the relaxation mechanism is found in the comprehensive work pursued by Bolkhovityanov *et al.* [37, 39, 40].

The two-step LT buffer with a Ge concentration of 70% served as a very useful template for strained Ge channel devices. The strained Ge channel formed on the two-step LT buffer exhibited much higher hole mobility than that formed on the one-step LT buffer [36], indicating greatly improved quality. Based on the two-step LT buffer, p-channel MOSFET was fabricated and very high hole effective mobility, exceeding that of bulk-Ge, was obtained [41]. It should be pointed out that the defect-induced SiGe layer is likely to become p-type in spite of no intentional doping, which forms a parallel path for hole conduction and causes leakage current.

7.4 Ion-implantation buffer

Introduction of the defective layer is a key concept in forming SiGe buffers with enhanced strain relaxation and low threading dislocation density. Instead of low-temperature growth, ion implantation is an alternative way to induce defects into the SiGe/Si heterostructures and has been employed to form SiGe virtual substrates. A fabrication technique utilizing ion implantation has been developed by many researchers and is currently becoming one of the best established and most feasible methods. This method has been applied to strained Si/Ge channel devices and has led to superior transport properties, i.e., mobility enhancements [42–46], low off-current [44] and suppression of self-heating effects [43, 46] owing to the small thickness of the SiGe layer.

The ion-implantation method has several advantages over the LT buffer method. The LT method is not suitable for a gas-source growth system such as CVD, where LT growth is difficult due to very low growth rates. In contrast, the growth system does not matter for the ion-implantation method, which includes no low-temperature growth. This compatibility with the CVD system makes the ion-implantation method very suitable for mass production. Moreover, defect states created by ion implantation are highly controllable compared to the LT buffer method, since they are determined only by conditions of the ex-situ ion-implantation process almost independently of growth conditions. Implantation-induced defects are quite stable against elevated temperatures, whereas one needs to ensure that point defects induced by LT growth remain during the SiGe growth at higher temperatures.

The fabrication procedure usually consists of ion implantation into a pseudomorphic SiGe layer and subsequent thermal annealing for strain relaxation promotion (Fig. 7.1(c)). Ion-implantation defects act as dislocation nucleation sites, which lower the misfit dislocation nucleation energy, and providing more misfit dislocations and lead to a higher degrees of relaxation compared to the growth without such implantation-induced defects. Since dislocations end at the defective regions, strain relaxation is expected to occur without increase in threading dislocation density.

Historically, the effects of ion implantation on strain relaxation were first investigated by Hull *et al.* in 1990 [47]. They demonstrated that the lattice damage introduced by B^+ or As^+ ion implantation can increase strain relaxation. Later on, several ion species, such as Ge^+ [48] and He^+ ions [49], were adopted and also found to be effective in accelerating dislocation generation and consequently increasing the degree of strain relaxation. These early studies revealed that the implantation-induced defects act as dislocation nucleation sources irrespective of implanted ion species. Threading dislocation densities were, however, still quite high, larger than $\sim 10^9$ cm^{-2}, which is unacceptable for high-mobility channel applications.

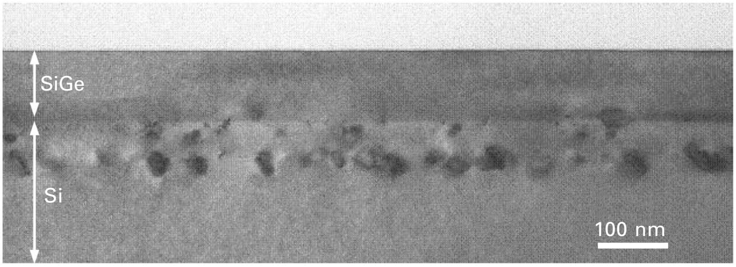
A comprehensive study to improve the quality of SiGe was pursued by Mantl and co-workers based on H^+ or He^+ implantation [50–53]. Pseudomorphic $Si_{1-x}Ge_x$ ($x = 15–30\%$) layers (100–250 nm) were grown on Si(100) substrates. The as-grown samples were subjected to H^+ or He^+ implantation at energies of 12–30 kV with doses of about $1–3 \times 10^{16}$ cm^{-2} and followed by thermal annealing at 600–1100°C. Implantation-induced defects were aimed to form about 50–100 nm below the SiGe/Si substrate heterointerface. Implantation doses were kept below 3×10^{16} cm^{-2} to avoid surface blistering. Note that those doses were about half of those doses used for the Smart-Cut process [54], an established technology for SOI wafer fabrication, where a Si thin overlayer is separated from the Si substrate via H-induced defects. During post-implantation annealing, high-density cavities (bubbles) filled with high-pressure H or He are formed underneath the heterointerface. These cavities generate dislocation loops which extend to the heterointerface and act as misfit segments resulting in strain relaxation.

Mantl and co-workers proposed a theoretical model to explain the defect-related relaxation mechanism as follows [51]. When the loop reaches a SiGe layer, it experiences an asymmetric force, and as a result, one side of the loop is pinned at the interface and forms misfit segments, whereas the other side is driven by the mismatch stress to the SiGe surface and forms a threading segment. Interaction of these threading dislocations with opposite Burgers vectors results in mutual annihilation of the dislocations, which effectively reduces threading dislocation density. Conditions required for effective relaxation and complete annihilation of threading dislocations were given as $\delta < 2.5d$, $\delta < h$, where δ is the average distance between the cavities, d is a loop diameter and h is a SiGe layer thickness. These conditions mean that higher density cavities, i.e. smaller δ, are necessary for formation of the SiGe buffers with the higher degree of relaxation, higher Ge concentrations in SiGe, and the smaller SiGe thickness. He^+ implantation was found to be more favorable for this requirement than H^+. TEM characterizations of a number of samples fabricated with various implantation and annealing conditions, they found that created He bubbles fulfilled these conditions (δ and d are several tens of nanometers). As a result, He^+ implantation successfully produced largely relaxed high-quality SiGe with Ge concentrations up to 30%, whereas H^+ implantation was revealed to be effective up to only about 20% Ge. With He^+ implantation at an energy of 18 keV with a dose of 2×10^{16} cm^{-2} and annealing at 750–850°C for 10 min, strain relaxation ratios of nearly 70% and threading dislocation densities below 5×10^7 cm^{-2} were achieved for 100 nm thick $Si_{0.7}Ge_{0.3}$ [53].

Although experimental observations were qualitatively in good agreement with the above-proposed model, it was pointed out that the conventional relaxation mechanism initiating from the surface or pre-existing dislocations simultaneously occurred in some conditions, which is likely to prevent efficient annihilation of threading dislocations. Moreover, highly complex cavity properties, especially their thermal evolution, made the relaxation behavior complicated. Also it was argued that the presence of He cavities around the heterointerface critically affects the threading dislocation density.

Cai *et al.* systematically investigated the effect of defects induced within a SiGe layer on threading dislocation [55]. It was found that the degree of strain relaxation depended strongly on the SiGe layer thickness but very weakly on the He^+ implantation conditions, i.e., the dose and energy. By contrast, threading dislocation density was revealed to be very sensitive to the implantation conditions, especially the amount of He residing within the SiGe layer. He^+ ions in the SiGe layer, the amount of which varies with the dose and energy, form cavities within the SiGe layer and impede the propagation of misfit dislocations, which results in short misfit dislocations and a high density of threading dislocations. It was concluded that creating high-density defects only underneath the SiGe/Si heterointerface, not within SiGe, is a crucial requirement to realize both large strain relaxation and low threading dislocation density. Based on their characterization of a large number of samples, they suggested that the lower limit of the threading dislocation density is 10^7 cm^{-2}.

For the purpose of avoiding defect formation in the SiGe layer, a pre-ion-implantation method was proposed [56, 57]. Ion implantation into a Si substrate is carried out prior to the SiGe growth. Relatively heavy ion species, such as Ar^+, Si^+ and Ge^+ [58], are favorable to form defects sufficiently near the Si surface, that is, at the Si/SiGe heterointerface. Otherwise, induced defects cannot act as dislocation nucleation sources. Since the implanted region becomes partially amorphous, the implanted Si substrate is next subjected to recrystallization annealing so that the epitaxial SiGe growth proceeds on the implanted Si. Even after this annealing procedure, defects remain in the Si substrate. Provided that the defects are buried slightly but sufficiently below the surface, the top surface is free from defects, allowing good epitaxial growth on the implanted Si substrate. Then, a SiGe layer is pseudomorphically grown. Here, growth at relatively low temperatures was found to be very important. At growth temperatures higher than 600°C, surface roughening due to elastic stress relief takes place and dislocation nucleation from the surface, not from the induced defects, dominates the relaxation mechanism [59]. On the other hand, by reducing the growth temperature down to 500°C, relaxation during growth was completely suppressed and a pseudomorphic layer is grown. Subsequently, post-growth annealing is performed and strain relaxation proceeds [57]. Figure 7.6 shows an XTEM image of 100 nm thick $Si_{0.8}Ge_{0.2}$ grown on the Si^+ implanted Si substrate. High-density dislocations are visible around the heterointerface, while almost no dislocation was found in the case without implantation. The relaxation ratio of the sample with implantation was more than 60% whereas that with



7.6 Cross-sectional TEM image of a 100 nm thick $Si_{0.8}Ge_{0.2}$ buffer layer grown on the Si^+-ion-implanted Si substrate.

no implantation was only several percent. This indicates that induced defects successfully act as dislocation nucleation sources and enhance the relaxation similarly to the He^+ implantation method. It was reported that states of defects induced by the heavy ions are quite different from those induced by light ions such as H^+ and He^+ [60]. It is considered that end-of-range defects play an important role in the relaxation mechanism for heavy ion cases [61].

An advantage of the pre-ion-implantation method is that the SiGe layer is free from implantation-induced defects, where misfit dislocation motion is not impeded. Moreover, high-density dislocations are uniformly nucleated, leading to a very homogeneous in-plane strain field distribution. A Raman mapping measurement revealed that a crosshatch-like strain distribution was completely eliminated owing to the implantation [62]. This uniformity was superior to those of conventional graded buffers and LT buffers. Moreover, surface smoothness was also much better than that obtained from other methods, with rms roughness less than 0.5 nm, which is comparable to the planarized surface.

Ion-implantation methods have been researched mainly for SiGe virtual substrates with Ge concentrations below 30%. Applying this method to higher Ge contents is the next challenge to widen the possibility of this method toward Ge channels and compound semiconductor materials. Growth of the pseudomorphic SiGe layer before past-growth annealing is an essential step to make implantation-induced defects more dominant relaxation sources and to suppress an occurrence of competitive conventional relaxation. In this sense, formation of SiGe with higher Ge content requires thinner SiGe thickness because the critical thickness is largely reduced with increase in mismatch strain of SiGe. However, sufficient thickness (more than several tens of nanometers) seems to be indispensable for the effective dislocation glide, which imposes an upper limit of Ge content around 30% even with metastable growth. This critical issue prevents successful application of this method to high Ge contents. Recently, Hoshi *et al.* have attempted fabrication of thin SiGe relaxed layers with Ge contents of nearly 50% [63], where the same process as the pre-implantation method was performed not directly on a Si substrate but on a conventional SiGe graded buffer layer with low Ge content. Since the mismatch strain between the upper and lower SiGe layers was kept sufficiently low, a high-quality SiGe with low threading dislocation density was attained, indicating the possible applicability of this method to high Ge contents. Several-step implantation and/or combination with other methods may be expected to extend the applicability.

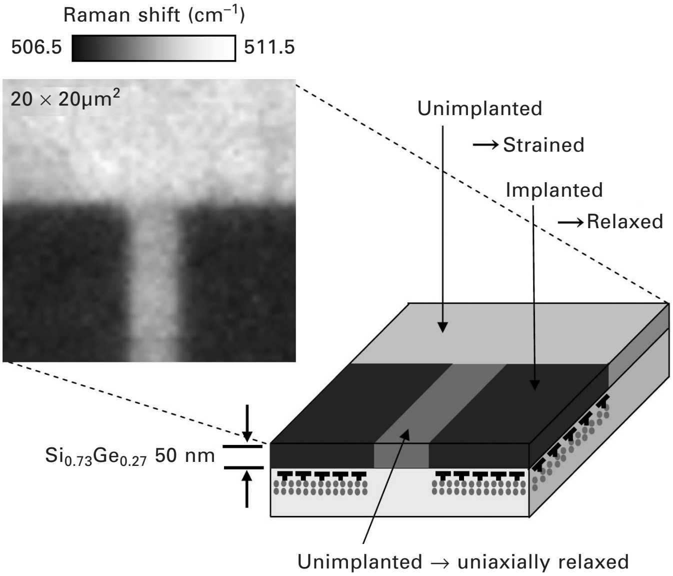
7.5 Other methods and future trends

Three major fabrication techniques for SiGe virtual substrates have been outlined. The graded buffer method is more advantageous for the purpose of reducing threading dislocation density. With surface planarization via CMP, almost ideal SiGe virtual substrates with negligible defect density are now available for a wide range of Ge concentrations. The only unsolved problem is its large thickness, which is indispensable for this method since the lower grading rate in general leads to a lower density of threading dislocations. Strain field fluctuation remaining at the SiGe surface is problematic as well. The other two techniques are recognized as methods that can modify the dislocation structure, that is, facilitate dislocation nucleation by means of intentionally introducing nucleation sources via low-temperature growth or ion bombardment. Although the issue of the thickness can be overcome by these methods, the attainable defect density is still not as low as that of the graded technique. Further improvements of these methods and/or developments of novel technologies, including combination of these techniques, are expected. Actually, a variety of techniques have been proposed in addition to the above three methods. In the last part of this chapter, some other methods and future prospects are presented.

One example is the use of a porous Si substrate as a defective region like LT-Si, where the so-called compliant effect was expected to suppress dislocation generation [64–66]. Concerning the compliant effect, growth on SOI substrates has been intensively studied, as described in the next chapter. By inserting a thin Ge layer prior to the SiGe growth, pure-edge dislocations other than 60° dislocations have been shown to be predominantly generated, contributing to reduction of mosaicity of the SiGe buffer layer [67]. More recently, a reverse graded method has been proposed [68] to offer high-quality SiGe virtual substrates with high Ge contents [69], in which direct growth of a thick Ge layer on a Si substrate was followed by SiGe growth with lower Ge concentrations. Since the Ge channel MOS device is indeed recognized as one of the most promising candidates for the next generation CMOS circuits, the role of SiGe virtual substrates with higher Ge compositions must became more and more important. From a device application point of view, fabrication technologies will need to become more sophisticated for realization of much thinner virtual substrates with much lower defect densities.

SiGe virtual substrates with different orientations such as (110) and (111) have also been studied by several groups [70, 71]. However, it was found to be more difficult to create them in the conventional graded manner unlike SiGe(100), since planar defects, such as stacking faults and micro-twins, are unavoidably induced and threading dislocations cannot be sufficiently reduced. Novel techniques are expected to be developed to overcome this problem.

As a modified ion-implantation method, a selective ion-implantation technique was proposed to create SiGe virtual substrates containing asymmetric (uniaxial) strain states [72, 73]. A concept of this method is to locally form a strained SiGe between relaxed SiGe regions on one planar wafer, where the strained SiGe can be anisotropically relaxed by shear stress coming from the neighboring relaxed SiGe regions. In-plane arbitrary control of strain states in SiGe can be attained by selective ion implantation since the relaxation of SiGe occurs only in implanted regions. Figure 7.7 shows a Raman mapping image of a SiGe grown on a Si substrate that was implanted with square patterns. As clearly seen in the image, relaxed and strained regions were simultaneously formed on a planar SiGe layer. It is also found that the unimplanted region between the implanted (and relaxed) SiGe exhibited a



7.7 Raman mapping image and sample schematic for the SiGe buffer layer grown on a Si substrate that was selectively Ar^+-ion-implanted with square patterns.

lower strain level than an unimplanted region outside the line pattern, indicating that the aforementioned stress-related relaxation takes place. More precise evaluation with XRD verified that highly asymmetric (uniaxial) strain states were realized with this method [72]. Recently, uniaxially strained channels have attracted increasing interest due to experimental findings that Si channels with uniaxial strain brought additional mobility enhancements exceeding those of the biaxial one. Since Si/Ge heterostructures have been manipulating only symmetric (biaxial) strain so far, the new technology, which can introduce asymmetric strain into SiGe virtual substrates, is expected to greatly widen the possibility of strain engineering based on Si/Ge heterostructures.

7.6 Sources of further information

E. A. Fitzgerald, M. T. Currie, S. B. Samavedam, T. A. Langdo, G. Taraschi, V. Yang, C. W. Leitz, and M. T. Bulsara, *Phys. Stat. Sol. (a)* **171**, 227 (1999).

R. Hull and J. C. Bean, *CRC Crit. Rev. Solid State Mater. Sci.* **17**, 507 (1992).

P. M. Mooney, *Mater. Sci. Eng.*, **R17**, 105 (1996).

P. M. Mooney and J. O. Chu, *Annu. Rev. Mater. Sci.* **30**, 335–362 (2000).

7.7 References

[1] C. G. Tuppen, C. J. Gibbings, and M. Hockly, *Mater. Res. Soc. Symp. Proc.* **220**, 187 (1991).

[2] E. A. Fitzgerald, Y. H. Xie, M. L. Green, D. Brasen, and A. R. Kortan, *Mater. Res. Soc. Symp. Proc.* **220**, 211 (1991).

[3] W. T. Pike, R. A. A. Kubiak, E. H. C. Parker, and T. E. Whall, *Mater. Res. Soc. Symp. Proc.* **220**, 223 (1991).

[4] F. K. LeGoues, B. S. Meyerson, and J. F. Morar, *Phys. Rev. Lett.* **66**, 2903 (1991).

[5] E. A. Fitzgerald, Y.-H. Xie, M. L. Green, D. Brasen, A. R. Kortan, J. Michel, Y.-J. Mii, and B. E. Weir, *Appl. Phys. Lett.* **59**, 811 (1991).

[6] Y.-J. Mii, Y.-H. Xie, E. A. Fitzgerald, D. Monroe, F. A. Thiel, B. E. Weir, and L. C. Feldman, *Appl. Phys. Lett.* **59**, 1611 (1991).

[7] E. A. Fitzgerald, Y.-H. Xie, D. Monroe, P. J. Silverman, J. M. Kuo, A. R. Kortan, F. A. Thiel, and B. E. Weir, *J. Vac. Sci. Technol. B* **10**, 1807 (1992).

[8] Y.-H. Xie, E. A. Fitzgerald, D. Monroe, P. J. Silverman, and G. P. Watson, *J. Appl. Phys.* **73**, 8364 (1993).

[9] K. Ismail, F. K. LeGoues, K. L. Saenger, M. Arafa, J. O. Chu, P. M. Mooney, and B. S. Meyerson, *Phys. Rev. Lett.* **73**, 3447 (1994).

[10] B. S. Meyerson, K. J. Uram, and F. K. LeGoues, *Appl. Phys. Lett.* **53**, 2555 (1988).

[11] F. K. LeGoues, B. S. Meyerson, J. F. Morar, and P. D. Kirchner, *J. Appl. Phys.* **71**, 4230 (1992).

[12] J. Tersoff and F. K. LeGoues, *Phys. Rev. Lett.* **72**, 3570 (1994).

[13] F. K. LeGoues, *Phys. Rev. Lett.* **72**, 876 (1994).

[14] P. M. Mooney, F. K. LeGoues, J. Tersoff, and J. O. Chu, *J. Appl. Phys.* **75**, 3968 (1994).

[15] M. A. Lutz, R. M. Feenstra, F. K. LeGoues, P. M. Mooney, and J. O. Chu, *Appl. Phys. Lett.* **66**, 724 (1995).

[16] T. Spila, P. Desjardins, J. D'Arcy-Gall, R. D. Twesten, and J. E. Greene, *J. Appl. Phys.* **93**, 1918 (2003).

[17] J. W. P. Hsu, E. A. Fitzgerald, Y.-H. Xie, P. J. Silverman, and M. J. Cardillo, *Appl. Phys. Lett.* **61**, 1293 (1992).

[18] S. B. Samavedam and E. A. Fitzgerald, *J. Appl. Phys.* **81**, 3108 (1997).

[19] C. W. Leitz, M. T. Currie, A. Y. Kim, J. Lai, E. Robbins, E. A. Fitzgerald, and M. T. Bulsara, *J. Appl. Phys.* **90**, 2730 (2001).

[20] M. T. Currie, S. B. Samavedam, T. A. Langdo, C. W. Leitz, and E. A. Fitzgerald, *Appl. Phys. Lett.* **72**, 1718 (1998).

[21] K. Sawano, K. Kawaguchi, T. Ueno, S. Koh, K. Nakagawa, and Y. Shiraki, *Mater. Sci. Eng.* **B89**, 406 (2002).

[22] K. Sawano, K. Kawaguchi, S. Koh, Y. Shiraki, Y. Hirose, T. Hattori, and K. Nakagawa, *J. Electrochem. Soc.* **150**, G376 (2003).

[23] M. T. Currie, C. W. Leitz, T. A. Langdo, G. Taraschi, E. A. Fitzgerald, and D. A. Antoniadis, *J. Vac. Sci. Technol. B* **19**, 2268 (2001).

[24] N. Sugii, D. Hisamoto, K. Washio, N. Yokoyama, and S. Kimura, *IEEE Trans. Electron Devices*, **49**, 2237 (2002).

[25] K. Sawano, Y. Hirose, S. Koh, K. Nakagawa, T. Hattori, and Y. Shiraki, *Appl. Phys. Lett.* **82**, 412 (2003).

[26] K. Sawano, Y. Abe, H. Satoh, K. Nakagawa, and Y. Shiraki, *Jpn. J. Appl. Phys.* **44**, L1320 (2005).

[27] K. Sawano, S. Koh, Y. Shiraki, N. Usami, and K. Nakagawa, *Appl. Phys. Lett.* **83**, 4339 (2003).

[28] K. Sawano, N. Usami, K. Arimoto, K. Nakagawa, and Y. Shiraki, *Jpn. J. Appl. Phys.* **44**, 8445 (2005).

[29] E. Kasper, K. Lyutovich, M. Bauer, and M. Oehme, *Thin Solid Films* **336**, 319 (1998).

[30] H. Chen, L. W. Guo, Q. Cui, Q. Hu, Q. Huang, and M. Zhou, *J. Appl. Phys.* **79**, 1167 (1996).

[31] J. H. Li, C. S. Peng, Y. Wu, D. Y. Dai, J. M. Zhou, and Z. H. Mai, *Appl. Phys. Lett.* **71**, 3132 (1997).

[32] K. K. Linder, F. C. Zhang, J.-S. Rieh, P. Bhattacharya, and D. Houghton, *Appl. Phys. Lett.* **70**, 3224 (1996).

[33] S. W. Lee, H. C. Chen, L. J. Chen, Y. H. Peng, C. H. Kuan, and H. H. Cheng, *J. Appl. Phys.* **92**, 6880 (2002).

[34] Y. H. Luo, J. Wan, R. L. Forrest, J. L. Liu, M. S. Goorsky, and K. L. Wang, *J. Appl. Phys.* **89**, 8279 (2001).

[35] A. P. Knights *et al.*, *J. Appl. Phys.* **89**, 76 (2001).

[36] T. Ueno, T. Irisawa, and Y. Shiraki, *J. Cryst. Growth* **227**, 761 (2001).

[37] Yu. B. Bolkhovityanov, A. S. Deryabin, A. K. Gutakovskii, M. A. Revenko, and L. V. Sokolov, *J. Appl. Phys.* **96**, 7665 (2004).

[38] C. S. Peng, Z. Y. Zhao, H. Chen, J. H. Li, Y. K. Li, L. W. Guo, D. Y. Dai, Q. Huang, J. M. Zhou, Y. H. Zhang, T. T. Sheng, and C. H. Tung, *Appl. Phys. Lett.* **72**, 3160 (1998).

[39] Yu. B. Bolkhovityanov, A. K. Gutakovskii, V. I. Mashanov, O. P. Pchelyakov, M. A. Revenko, and L. V. Sokolov, *Thin Solid Films* **392**, 98 (2001).

[40] Yu. B. Bolkhovityanov, A. K. Gutakovskii, V. I. Mashanov, O. P. Pchelyakov, M. A. Revenko, and L. V. Sokolov, *J. Appl. Phys.* **91**, 4710 (2002).

[41] T. Irisawa, S. Tokumitsu, T. Hattori, K. Nakagawa, S. Koh, and Y. Shiraki, *Appl. Phys. Lett.* **81**, 847 (2002).

[42] H.-J. Herzog, T. Hackbarth, U. Seiler, U. König, M. Luysberg, B. Holländer, and S. Mantl, *IEEE Electron Device Lett.* **23**, 485 (2002).

[43] T. Hackbarth, H.-J. Herzog, K. H. Hieber, U. König, S. Mantl, B. Holländer, S. Lenk, H. von Känel, M. Enciso, F. Aniel, and L. Giguerre, *Solid-State Electronics* **48**, 1921 (2004).

[44] P. M. Mooney, K. Rim, S. H. Christiansen, K. K. Chan, J. O. Chu, J. Cai, H. Chen, J. L. Jordan-Sweet, Y. Y. Yang, and D. C. Boyd, *Solid-State Electronics* **49**, 1669 (2005).

[45] K. Sawano, A. Fukumoto, Y. Hoshi, Y. Shiraki, J. Yamanaka, and K. Nakagawa, *Appl. Phys. Lett.* **90**, 202101 (2007).

[46] L. Yan, S. H. Olsen, E. Escobedo-Cousin, and A. G. O'Neill, *J. Appl. Phys.* **103**, 094508 (2008).

[47] R. Hull, J. C. Bean, J. M. Bonar, G. S. Higashi, K. T. Short, H. Temkin, and A. E. White, *Appl. Phys. Lett.* **56**, 2445 (1990).

[48] G. P. Watson, E. A. Fitzgerald, Y.-H. Xie, P. J. Silverman, A. E. White, and K. T. Short, *Appl. Phys. Lett.* **63**, 746 (1993).

[49] D. M. Follstaedt, S. M. Myers, and S. R. Lee, *Appl. Phys. Lett.* **69**, 2059 (1996).

[50] B. Holländer, S. Mantl, R. Liedtke, S. Mesters, H.-J. Herzog, H. Kibbel, and T. Hackbarth, *Nucl. Instrum. Methods B* **148**, 200 (1999).

[51] H. Trinkaus, B. Holländer, S. Rongen, S. Mantl, H.-J. Herzog, J. Kuchenbecher, and T. Hackbarth, *Appl. Phys. Lett.* **76**, 3522 (2000).

[52] B. Holländer, S. Lenk, S. Mantl, H. Trinkaus, D. Kirch, M. Luysberg, T. Hackbarth, H.-J. Herzog, and P. F. P. Fichtner, *Nucl. Instrum. Methods B* **175**, 357 (2001).

[53] M. Luysberg, D. Kirch, H. Trinkaus, B. Holländer, S. Lenk, S. Mantl, H.-J. Herzog, T. Hackbarth, and P. F. P. Fichtner, *J. Appl. Phys.* **92**, 4290 (2002).

[54] M. Bruel, *Nucl. Instrum Methods B* **108**, 313 (1996).

[55] J. Cai, P. M. Mooney, S. H. Christiansen, H. Chen, J. O. Chu, and J. A. Ott, *J. Appl. Phys.* **95**, 5347 (2004).

[56] K. Sawano, Y. Hirose, Y. Ozawa, S. Koh, J. Yamanaka, K. Nakagawa, T. Hattori, and Y. Shiraki, *Jpn. J. Appl. Phys.* **42**, 735 (2003).

[57] K. Sawano, S. Koh, Y. Shiraki, Y. Ozawa, T. Hattori, J. Yamanaka, K. Suzuki, K. Arimoto, K. Nakagawa, and N. Usami, *Appl. Phys. Lett.* **85**, 2514 (2004).

[58] K. Sawano, A. Fukumoto, Y. Hoshi, J. Yamanaka, K. Nakagawa, and Y. Shiraki, *Thin Solid Films* **517**, 87 (2008).

[59] K. Sawano, Y. Hirose, S. Koh, K. Nakagawa, T. Hattori, and Y. Shiraki, *Appl. Surf. Sci.* **224**, 99 (2004).

[60] D. Buca, R. A. Minamisawa, H. Trinkaus, B. Holländer, S. Mantl, R. Loo, and M. Caymax, *J. Appl. Phys.* **105**, 114905 (2009).

[61] Y. Hoshi, K. Sawano, A. Yamada, K. Arimoto, N. Usami, K. Nakagawa, and Y. Shiraki, *Thin Solid Films* **518**, S162 (2010).

[62] K. Sawano, Y. Ozawa, A. Fukumoto, Y. Shiraki, J. Yamanaka, K. Suzuki, K. Arimoto, K. Nakagawa, and N. Usami, *Jpn. J. Appl. Phys.* **44**, L1316 (2005).

[63] Y. Hoshi, K. Sawano, Y. Hiraoka, Y. Satoh, Y. Ogawa, A. Yamada, N. Usami, K. Nakagawa, and Y. Shiraki, *Appl. Phys. Express* **1**, 081401 (2008).

[64] S. I. Romanov, V. I. Mashanov, L. V. Sokolov, A. Gutakovskii, and O. P. Pchelyakov, *Appl. Phys. Lett.* **75**, 4118 (1999).

[65] Y.-H. Xie and J. C. Bean, *J. Vac. Sci. Technol. B* **8**, 227 (1990).

[66] N. Usami, K. Kutsukake, K. Nakajima, S. Amtablian, A. Fave, and M. Lemiti, *Appl. Phys. Lett.* **90**, 031915 (2007).

[67] A. Sakai, N. Taoka, O. Nakatsuka, S. Zaima, and Y. Yasuda, *Appl. Phys. Lett.* **86**, 221916 (2005).

[68] L. H. Wong, J. P. Liu, F. Romanato, C. C. Wong, and Y. L. Foo, *Appl. Phys. Lett.* **90**, 061913 (2007).

[69] V. A. Shah, A. Dobbie, M. Myronov, D. J. F. Fulgoni, L. J. Nash, and D. R. Leadley, *Appl. Phys. Lett.* **93**, 192103 (2008).

[70] M. L. Lee, D. A. Antoniadis, and E. A. Fitzgerald, *Thin Solid Films* **508**, 136 (2006).

[71] V. Destefanis, J. M. Hartmann, A. Abbadie, A. M. Papon, and T. Billon, *J. Cryst. Growth* **311**, 1070 (2009).

[72] K. Sawano, Y. Hoshi, A. Yamada, Y. Hiraoka, N. Usami, K. Arimoto, K. Nakagawa, and Y. Shiraki, *Appl. Phys. Express* **1**, 121401 (2008).

[73] K. Sawano, Y. Hoshi, Y. Hiraoka, N. Usami, K. Nakagawa, and Y. Shiraki, *J. Cryst. Growth* **311**, 806 (2009).

8

Formation of silicon–germanium on insulator (SGOI) substrates

N. SUGIYAMA and T. TEZUKA, Toshiba Corporation,
Japan

Abstract: A new technology, called the Ge condensation method, that realizes a thinner SiGe layer with higher Ge content by the high-temperature oxidation of silicon–germanium on insulator (SGOI) substrate, is proposed. The mechanism of the Ge condensation method and the efforts to achieve a high-quality SGOI layer suitable for high-performance strained-Si MOSFET are explained, and the Ge condensation method is compared with other technologies for preparing SGOI and strained Si on insulator structure.

Key words: SiGe on insulator, strained silicon, germanium condensation, strain relaxation, dislocation.

8.1 Introduction: demand for virtual substrate and (Si)Ge on insulator (SGOI)

Tremendous efforts have been made in material science to enhance the performance of semiconductor devices. In terms of the metal-oxide-semiconductor field-effect transistor (MOSFET), adoption of new materials such as strained Si and SiGe in the channel region is one of the most attractive approaches because a higher drive current can be obtained by higher mobility in the channel layers [1–3].

In the case of a strained Si channel, it is necessary to prepare a template crystal substrate that has a larger in-plan lattice spacing than the Si crystal. For this purpose, thick SiGe crystal on Si substrate, which is lattice-relaxed by introducing many dislocations, is one of the materials suitable for virtual substrates [4–9]. On the other hand, in view of the desirability of reducing junction capacitance, a thin semiconductor layer on the insulator structure is another requirement for high-performance MOSFET. For this purpose, epitaxial growth of the SiGe layer on thin SOI (Si on insulator) substrate has been examined with a view to obtaining a relaxed SiGe layer with a relatively thinner layer [10].

Another way of obtaining a thin SiGe layer relaxed on the insulator layer has been proposed, namely the application of SIMOX technology, which consists of oxygen ion implantation and high-temperature annealing, on the thick SiGe layer grown on Si substrate. The first report of a strained-Si layer on a thin SiGe layer laminated on the silicon dioxide layer, called SiGe on insulator (SGOI), was published in 1999 [11]. Subsequently, demonstration of high-performance MOSFET verified that the SGOI structure is suitable for MOSFET devices from the viewpoint of low capacitance and higher channel mobility [12].

However, it has been revealed that the SIMOX technology cannot be applied on a SiGe layer with Ge content higher than 15%, which has been mandatory for higher strain in the Si layer, owing to the limitation of annealing temperature imposed by the melting point of SiGe crystal [13]. In order to overcome this limitation, a new technology, called the Ge condensation method, that realizes a thinner SiGe layer with higher Ge content by the high-temperature oxidation of SGOI, has been proposed [14, 15]. In this chapter, the mechanism of the Ge condensation method and the efforts to achieve a high-quality SGOI layer suitable for high-performance strained-Si MOSFET are explained, and the Ge condensation method is compared with other technologies for preparing SGOI and strained Si on insulator structure.

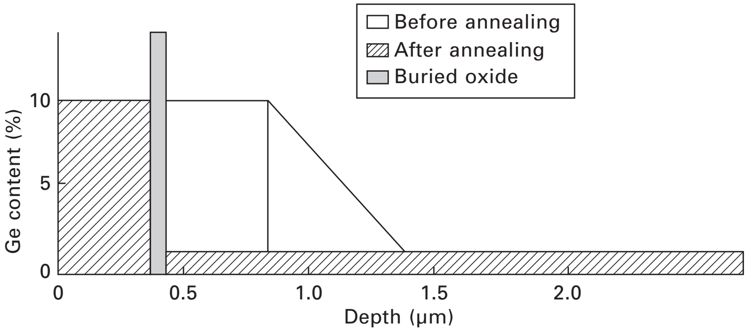
8.2 Formation of (Si)Ge on insulator (SGOI) by the Ge condensation method

8.2.1 Fundamentals of the Ge condensation method

Limit of SIMOX technology and introduction of the Ge condensation method

Before discussing the mechanism of Ge condensation, it is necessary to briefly explain the limitation of the SIMOX process for SGOI formation. SIMOX technology consists of two processes: oxygen ion implantation onto the SiGe crystal layer, and high-temperature annealing to form a flat buried oxide layer. Typically, a temperature as high as 1350°C is necessary for the formation of a buried oxide layer in Si crystal. On the other hand, since the melting point of SiGe crystal decreases with increasing Ge content, it is also necessary to lower the annealing temperature for the formation of the buried oxide (BOX) layer among SiGe [16]. For example, when the sample is annealed at 1350°C, a uniform and flat oxide layer can be obtained. However, it has been experimentally revealed that the flatness of the buried oxide layer is degraded and it becomes wavy at 1320°C annealing. And at 1290°C, a continuous oxide layer cannot be obtained. This means that the temperature required for the flat buried oxide formation does not decrease with the Ge content, though the melting point decreases.

Next, the unique behavior of Ge atoms during the SIMOX process that leads to the primitive idea of the Ge condensation method is explained. In Fig. 8.1, the Ge profile before and after the SIMOX process is schematically

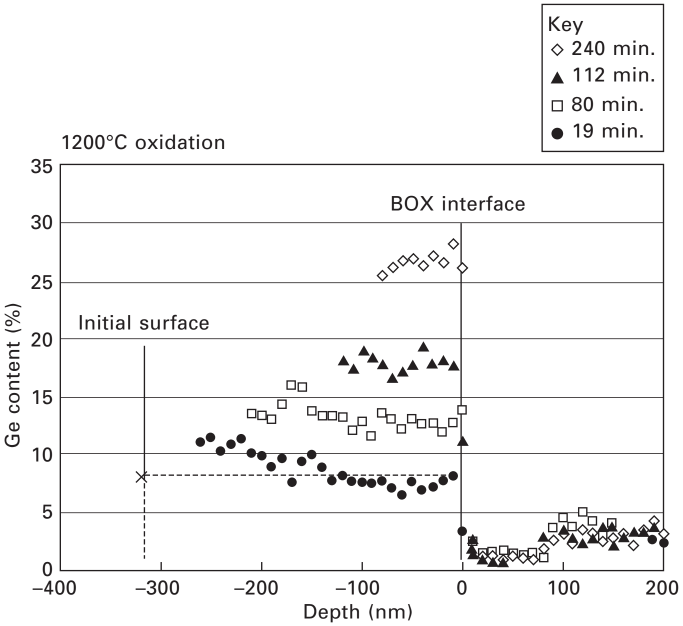


8.1 Schematic illustration of Ge profiles before and after the SIMOX process.

illustrated. Here, the Ge content in the SGOI layer on the buried oxide is different from that in the layer under the buried oxide, and the Ge content in buried oxide is quite low. SIMS analysis has revealed that the Ge content in the buried oxide is around 10^{18} cm^{-3}. These results include two important points that lead to the idea of a new process, named the Ge condensation method. The first point is that only a few Ge atoms exist in the buried oxide layer, although the oxygen ions have been implanted into the SiGe layer to form the buried oxide. This means that the Ge atoms are rejected from the oxide layer during the formation of the oxide by high-temperature annealing. These rejected Ge atoms are thought to be accumulated at the interface of oxide and SiGe layer [17], and they diffuse into the SiGe crystal layer in the case of high-temperature oxidation. The second point is that the Ge content in the SiGe layer upon the buried oxide layer remains at 10%, whereas the Ge atoms in the layer under the buried oxide diffuse into the substrate. This means that the buried oxide layer acts as a barrier against the Ge diffusion. Consideration of these two characteristics of the SIMOX process for SGOI formation leads to the following idea: the SiGe layer sandwiched by the thermal oxide layer and the buried oxide layer becomes thinner with the increase of the Ge content through the thermal oxidation process, because the Ge atoms rejected from the thermal oxide layer are incorporated into the SiGe layer but cannot diffuse over the buried oxide barrier during the oxidation.

Verification of the Ge condensation mechanism

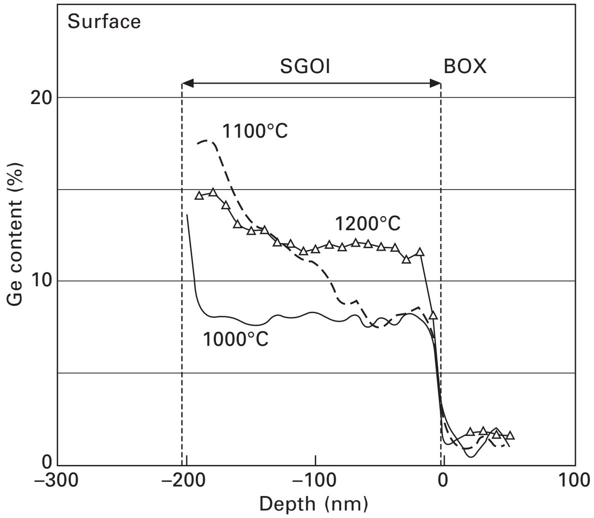
In order to prove the above condensation mechanism, a SGOI layer has been oxidized in an oxygen atmosphere. Figure 8.2 indicates the variation of the Ge profile obtained by sputtering Auger analysis depending on the oxidation time at 1200°C. It is confirmed that the Ge content increases according to



8.2 Variation of depth profiles of Ge atoms in the SGOI structure formed by the Ge condensation method with various oxidation times.

the reduction of the SGOI thickness by oxidation at the high temperature. Here, it is confirmed that the total amount of Ge atoms in the SGOI layer remains constant during oxidation. This means that almost the total amount of Ge atoms is conserved because the Ge atoms scarcely evaporate from the surface and scarcely diffuse over the oxide barrier during the oxidation.

On the other hand, it is found that the piling up of Ge atoms at the interface increases with a decrease in the oxidation temperature, as shown in Fig. 8.3. The dependence of the variation of the Ge depth profile in the SGOI layer on the oxidation temperature is explained as follows. At first, the Ge atoms are not incorporated into the oxidized layer during the high-temperature oxidation of the SiGe layer, but accumulate at the SiGe/thermal oxide interface. On the other hand, the accumulated Ge atoms diffuse toward the substrate. Thus, the depth profile of Ge atoms near the thermal oxide interface is explained by the competitive process between the accumulation of Ge atoms at the oxide interface and the diffusion of Ge atoms toward the substrate. For example, abrupt profiles of Ge atoms are formed when the accumulation rate of the Ge atoms is much larger than the diffusion rate of Ge atoms toward the substrate. Then the Ge profile changes according to the temperature, because the temperature dependence of the accumulation rate and that of the diffusion rate are different.



8.3 Temperature dependence of the Ge profile in the SGOI structure formed by the Ge condensation process.

The temperature dependence of the accumulation rate of Ge atoms can be estimated from that of the oxidation rate, because the accumulation of Ge atoms is caused by the rejection of Ge atoms during the oxidization of the SiGe layer. Here, since the oxidation rate is regulated by the transport of oxygen among SiO_2 under the diffusion-limited conditions, it is reasonable to consider the activation energy of diffusion of oxygen in SiO_2 to be the key factor for the temperature dependence of the accumulation rate of Ge atoms at the interface. The activation energy of the oxygen diffusion in SiO_2 is reported to be 1.17 eV [18]. On the other hand, the activation energy of the Ge diffusion in the Si layer is reported to be 5.8 eV [19]. The fact that the activation energy of the Ge diffusion in Si is much higher than that of the oxygen diffusion in SiO_2 means that the temperature dependence of the Ge diffusion is stronger than that of the oxygen diffusion in SiO_2, which represents the accumulation rate of Ge. As a result, the Ge diffusion is more enhanced at higher temperature than the accumulation of Ge atoms at the SiO_2/SiGe interface, which leads to the flat Ge profile. On the other hand, at lower temperatures, the diffusion of Ge atoms near the interface is suppressed and the abrupt profile is formed. It should be noted, however, that there is an upper limit to the oxidation temperature, which comes from the decrease of the melting points of SiGe alloys with an increase in Ge content, as explained above. The melting point decreases to 1230°C at a Ge content of 30%.

8.2.2 Strain relaxation during Ge condensation

Continuous SGOI layer having no boundary

In the previous section, the simple mechanism of Ge condensation from the SGOI structure formed by the SIMOX process was explained. In this section, the relaxation mechanism of strained SiGe crystal during the Ge condensation process is discussed, referring to the example of an advanced condensation process, in which the SiGe epitaxial layer grown on a SOI substrate is condensed by oxidation [20]. In this process, the SiGe layer thinner than the critical thickness, which is fully strained and has no dislocations, is grown on a SOI substrate, and then the layer is oxidized at high temperature. During the oxidation, because of the diffusion of Ge atoms, the initial SOI layer disappears immediately and a uniform Ge profile is obtained. In this condensation process, the Ge content of the SiGe layer becomes higher in the same manner as in the case of the condensation from the SGOI formed by the SIMOX method. The in-plane lattice spacing of SiGe crystal, which is the same as that of Si crystal just after epitaxial growth, expands toward the intrinsic lattice spacing of fully relaxed SiGe crystal, as the condensation proceeds. Thus, the relaxation mechanism during the condensation process is important in this method.

Next, the relaxation process during the Ge condensation is discussed. Continuous SGOI layers are known to relax the internal strain during the condensation process when the layers are sufficiently thicker than a critical value. On the other hand, no relaxation has been observed for thinner SGOI layers. This behavior is similar to that of a SiGe layer on a Si substrate in the case of introducing misfit dislocations. Since formation of dislocations has been observed in the relaxed SGOI layers as discussed in Section 8.2.3, the relaxation mechanism for SGOI layers is also considered to be a plastic deformation by dislocations. Relaxation is promoted by dislocation generation along the {111} slip plane due to the shared component of the internal stress along the plane.

The relaxation behavior has been found to be dependent on the thickness of the SiGe layer. The relation between residual strain, ε, in the SGOI layer that is not fully relaxed and its in-plane lattice spacing of $a'(x)$ is described as

$$\varepsilon = [a(x) - a'(x)]/a(x) \tag{8.1}$$

Here $a(x)$ is the intrinsic lattice constant of SiGe crystal having a Ge content of x. A degree of relaxation, a relaxation ratio R, is defined as follows:

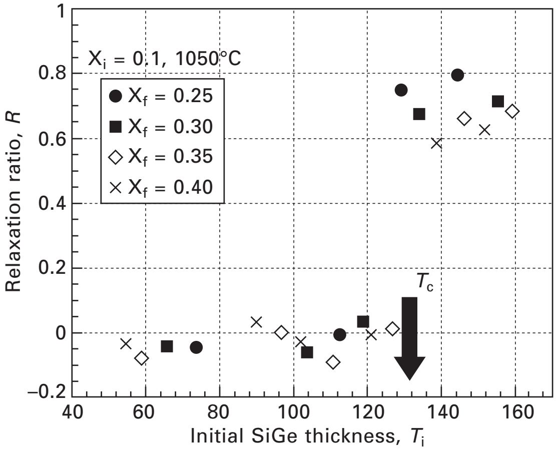
$$R = \frac{a'(x) - a(0)}{a(x) - a(0)} \tag{8.2}$$

Here, $a(0)$ is the lattice constant of Si crystal, and R values of 1 and 0 represent the SiGe layer having the same lattice constant as the intrinsic values for $Si_{1-x}Ge_x$ and Si, respectively.

The dependence of the relaxation ratio on the initial thickness of the SiGe layer has been experimentally evaluated. In order to evaluate the strain, ε, Raman spectroscopy is conventionally used. The relation between the strain, ε, and the Raman peak shift, $\Delta\omega$, is represented as follows:

$$\Delta\omega = \alpha x - \kappa\varepsilon(x) \tag{8.3}$$

Here, α and κ are the constantf coefficients, for which values in the range of 62–68 and 720–790, respectively, have been reported [21, 22]. Therefore, strain in the SiGe crystal can be measured by Raman spectroscopy when the Ge content is fixed. In experiment, samples having a SiGe layer of 50–190 nm with fixed Ge content ($X = 0.1$), which are grown on the SOI layer of 30 nm thickness, are prepared and then oxidized under an oxygen atmosphere at 1200°C. As a result, SGOI layers having Ge content of 0.1–0.6 are obtained. Figure 8.4 shows the relaxation ratio as a function of initial SiGe thickness for each final Ge content. This graph indicates the existence of the critical value for the initial SiGe thickness to form the relaxed SGOI layer. The critical value for the initial SiGe layer is estimated to be 130 nm, which does not depend on the Ge content of the final SGOI layer. The fact that the SGOI layer can be relaxed when the initial SiGe layer is thicker than the critical value implies that a plastic deformation of the initial SiGe layer initiates the relaxation process during high-temperature annealing. Details of the relaxation mechanism are discussed in Section 8.2.3.



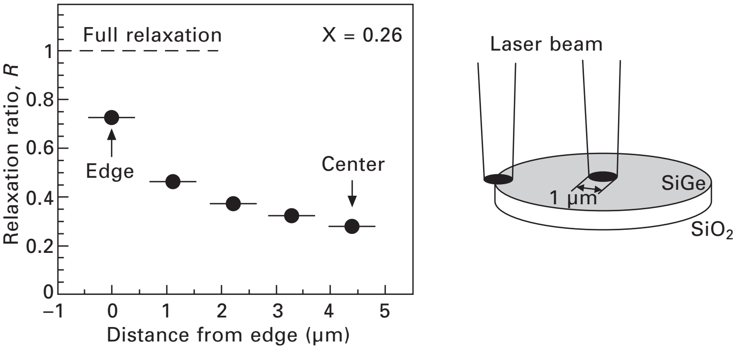
8.4 Relation between the relaxation ratio of the SGOI structure formed by the Ge condensation process and the initial SiGe thickness grown on SOI substrate.

SGOI layer shaped as a mesa island

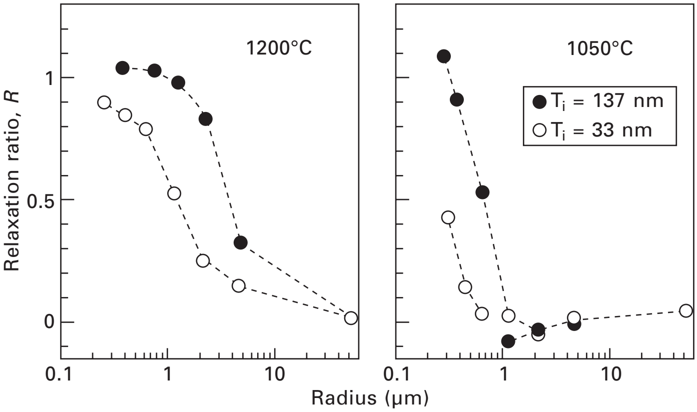
Next, a refined Ge condensation technique combined with mesa isolation in order to enhance the relaxation is introduced and its relaxation mechanism is discussed [23, 24]. In this method, because the boundary of a small mesa island can be the free edge for strain relaxation, an elastic deformation from the edge with a viscous deformation of the BOX layer is a major origin of the relaxation. The fabrication processes of the SGOI mesas are follows. First, pseudomorphic SiGe layers of thickness T_i and Ge content of 10% are epitaxially grown on SOI layers 30 nm thick. Next, circular mesas of the SiGe and Si layers are defined by chemical dry etching. Finally, these mesas are oxidized in a pure oxygen atmosphere at 1200 and 1050°C. The SGOI mesas are relaxed after the oxidation. Here, two wafers with T_i of 137 and 33 nm are used in the experiments.

Figure 8.5 shows the strain distribution along the radial direction for the SGOI mesa with $r = 4.4$ µm after oxidation at 1200°C. The plot indicates that the strain is minimal around the boundary and increases toward the center. In addition, the lattice relaxes partially at the center of the mesa. This result indicates that the lattice relaxation proceeds from the edge region toward the center region, because the compressive strain is effectively released and then relaxation is achieved at the edge region. Figure 8.6 shows plots of the relaxation ratio, *R*, at the center of the mesa, as functions of the mesa radius for thicker and thinner mesas, and for higher and lower oxidation temperatures. As shown in Fig. 8.6, the trend that *R* increases with increasing thickness and with decreasing radius is observed. It is also confirmed that *R* increases with the oxidation temperature. In the case of the starting SiGe layer thickness of 137 nm with 1200°C oxidation, a relaxation ratio of more than 80% is observed in the sample having a radius of less than 2 µm.

These dependences of the relaxation ratio on the SiGe layer thickness and the oxidation temperature can be explained by a simple relaxation model of



8.5 Strain distribution along the radial direction for the SGOI mesa with $r = 4.4$ µm after oxidation at 1200°C.



8.6 Relation between the relaxation ratio at the center of the mesa and the mesa radius for thicker and thinner mesas, and for higher and lower oxidation temperatures.

the mesas as follows. The driving force of the lattice relaxation is the shear stress at the interface between the SGOI and the buried oxide (BOX) layers, which originates in the compressive strain in the SGOI layer. Here, the shear stress is assumed to be uniform in the direction vertical to the substrate, and is described by a function of the distance from the center, d, for simplicity. Then, the shear stress has its maximum value at the mesa boundary, $d = r$, and its minimum at the center, $d = 0$, because the boundary is considered to be a free edge [25]. Before the oxidation, the shear stress cannot be released by the rigid SGOI/BOX interface at low temperatures.

The high temperature during the oxidation enables lateral expansion of the SGOI via slip at the interface and/or viscous flow of the BOX layer within the region of $r > d > r\ L_c$, where the shear stress overcomes the critical value defined by the sticking strength between the SGOI layer and the BOX layer [26]. The critical distance from the boundary, L_c, is considered to increase with the thickness of the SGOI layer as well as with the oxidation temperature, because the stress is roughly proportional to the thickness and the bonding strength at the SGOI/BOX interface becomes weaker at higher temperatures. When $L_c > r$, the entire mesa structure is relaxed. There is a good match between the temperature and thickness dependencies of R found in Fig. 8.6 and the above relaxation model of stress release via the slip at the SGOI/BOX interface. In addition, other experiments have revealed that the threading dislocation density in SGOI mesa of $r < 2$ μm is much lower than that in the continuous SGOI layer having no boundary. Correspondingly, the cross-hatched roughness originating from the piling up of misfit dislocations could not be observed on the mesa SGOI surface. These results also indicate that the strain relaxation during the Ge condensation of the mesa SGOI can be brought about by the lateral expansion of the SGOI via an elastic deformation rather than a plastic deformation of the SGOI layer. The relaxation behavior in Fig. 8.6 shows a monotonic decrease in the relaxation ratio and is almost zero for larger mesa SGOI layers of $r \sim 10$ μm. On the other hand, Fig. 8.4 indicates that the continuous SGOI layer with the same initial SiGe thickness of the thicker mesas in Fig. 8.6 is relaxed after the process. The discrepancy in the relaxation behavior between the mesa-isolated and the continuous layers can be explained in terms of a limited line density of misfit dislocations that defines the degree of relaxation in the case of a plastic deformation. Since the line density is roughly proportional to the area of the strained layer if the dislocations can extend between the boundaries, the mesas of $r \sim 10$ μm are considered to have significantly smaller dislocation densities than that for a continuous SGOI layer, resulting in a negligibly small relaxation ratio.

8.2.3 Mechanism of dislocation formation and reduction of dislocation

Evaluation technique for dislocations in SGOI

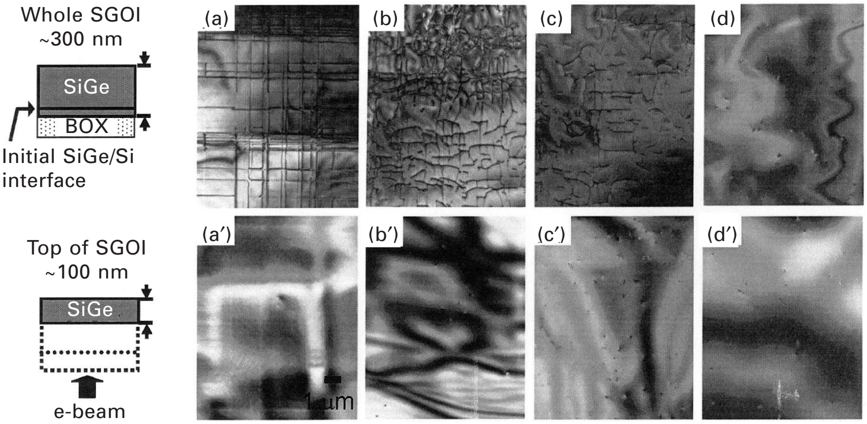
Before discussing the behavior of dislocations, it is necessary to consider how to observe dislocations in the thin SGOI layer. Conventionally, two methods are used for the evaluation of dislocation density in Si substrates: plan-view transmission electron microscopy (TEM) and etch-pit-density (EPD) observation after selective etching of dislocations. The plan-view TEM observation is the most reliable method. However, since this method requires large magnification, the detection limit of the dislocation density is high. For example, even if the sample has the threading dislocation (TD) density of 10^8 cm^{-2}, few dislocations are observed in an image (having a size of 30 cm × 20 cm) obtained by TEM observation with a magnification of 10^5 times, which is a typical magnification of plan-view TEM analysis for dislocation evaluation. In other words, no dislocations can be found in the TEM observation area for low dislocation densities of less than 10^7 cm^{-2}. EPD observation after a selective chemical etching method such as Secco etching is also an established technique for evaluating dislocation density in Si substrates. In this method, however, sufficient etching depth to form large pits, which can be easily observed using a microscope, is required. This means that, when the film is thin, the formation of clear etch pits is difficult.

In order to enlarge traces of small etch pits on thin SOI layers, an advanced method called enhanced Secco etching has been proposed [27]. In this method, the small pits formed by Secco etching are decorated with large disk-shaped holes formed inside the buried oxide by subsequent HF etching. These large holes make it possible to easily count the number of dislocations by using an optical microscope. In this method, if the dislocation is decollated with a hole having a size of 1 μm, several pits can be observed in the image obtained by an optical microscope with a magnification of 10^3 times, even if the dislocation density is as low as 10^4 cm^{-2}. Applicability of this technique to the SGOI structure has also been confirmed [28].

Incidence of threading dislocations during Ge condensation

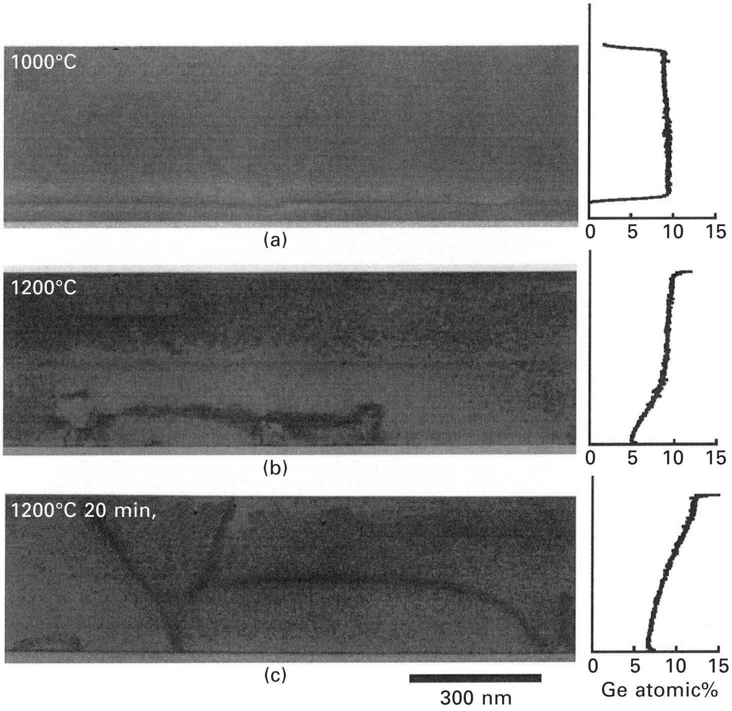
Next, the mechanism of the generation of dislocations during Ge condensation is discussed. It has been revealed, by the following experiment, that the behavior of misfit dislocations (MD) at the interface of the SiGe epitaxial layer and SOI plays an important role with regard to the incidence of TDs during the initial stage of the Ge condensation process. In an experiment, samples having a 330 nm SiGe ($X = 0.1$) epitaxial layer on SOI substrate were prepared for the initial structure for Ge condensation. The samples were oxidized in an oxygen atmosphere in the thermal sequences indicated as follows. Sequences A–D correspond to ramping-up processes to temperatures of (A) 900°C, (B) 1000°C, (C) 1100°C and (D) 1200°C, respectively. Sequences E and F correspond to condensation processes by oxidation at 1200°C for (E) 20 min and (F) 300 min, respectively.

Dislocation behavior during the condensation process was investigated by plan-view (Fig. 8.7) and cross-sectional (Fig. 8.8) TEM measurements. In Fig. 8.7, plan-view TEM images for whole SGOI layers (a–d) and top regions of SGOI layers (a′–d′) are shown. Here, the former involves initial SiGe/SOI interfacial regions, whereas the latter does not. A clear dislocation network



B. 1000°C D. 1200°C E. 1200°C 20 min F. 1200°C 300 min

8.7 Plan-view TEM images for whole SGOI layers (a–d) and top regions of SGOI layers (a′–d′) formed by the various thermal sequences of the Ge condensation method.



8.8 Cross-sectional TEM images of SGOI layers formed by the various thermal sequences of the Ge condensation method.

with the orientations of <110> is observed after ramping up to 1000°C, as indicated in Fig. 8.7(a). The dislocations are found to be located at the initial SiGe/SOI interface from Fig. 8.8(a). Therefore, these dislocations are considered to be misfit dislocations (MDs) formed by plastic relaxation of the SiGe layer. The density of threading arms of these dislocations, which might be terminated at the surface, is so low that no threading dislocations (TDs) are detected in the TEM image of Fig. 8.7(a′). As the condensation proceeds, it is found from Figs 8.7(b)–(d) and (b′)–(d′) that MDs fragment and then disappear, whereas TDs newly appear, respectively. The evolution of dislocations from MDs to TDs is shown in Figs 8.8(a)–(c). First, Fig. 8.8(a) shows that long MDs extending over several μm form at the SiGe/SOI interface during the ramping-up process. Then, in Fig. 8.8(b) it is seen that the MDs are partially absorbed into the SGOI/BOX interface, resulting in dislocation half-loops that terminate at the BOX interface upon completion of the ramping up. The half-loops, which are fragments of initial MDs, rise toward the surface in the oxidation process. After the loop reaches the surface, TDs are left in the SGOI layer (c). As a result, it is revealed that TD density is drastically increased after the ramping-up process to 1100°C

(C–E), and then saturates after a sufficiently long oxidation time of 300 min (F).

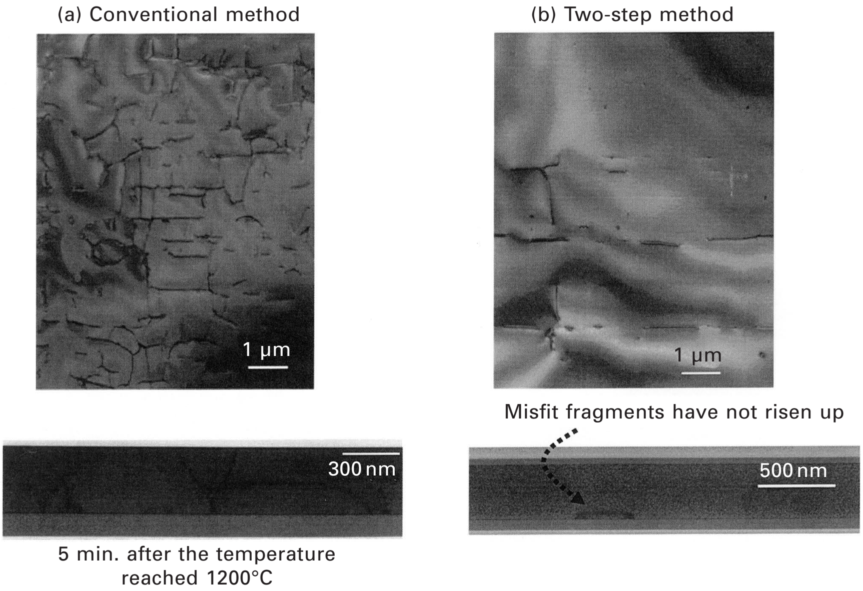
In the descending section, the dislocation motion is considered to be ascribable to one of three mechanisms or a combination of them: the repulsive interaction between adjacent or crossing dislocations [29], the image force to push dislocations to BOX interfaces [30] and the compressive-strain accumulation in underlying SOI layers due to interdiffusion of Si and Ge. The first two mechanisms have been described as driving forces for the descending motion of MDs by a glide mechanism in systems of epitaxial SiGe layers on SOI substrates [30–32]. The last-named mechanism may be specific to the Ge-condensation process because the oxidation temperature is so high that the initial SOI layer turns into a strained SiGe layer owing to interdiffusion of Si and Ge. The accumulated stress is expected to be a driving force of the descending motion of dislocations into the strained SiGe layer by a glide mechanism, since it can reduce the strain energy in the layer below the initial SiGe/SOI interface. When the dislocations have reached the BOX interface, it is considered that dislocation images become invisible in TEM images since the line tension of the dislocation cores is relaxed via the atomic reorientation of the amorphous BOX layer [31].

Proposal of the two-step Ge condensation method

In contrast to the case with the descending motion, the rising motion of MDs and the subsequent TD formation have not been reported in previous works concerning SiGe/SOI systems in non-oxidizing atmosphere [30–33]. Thus, the rising motion can be interpreted as an oxidation- or a condensation-induced phenomenon. A possible mechanism for the origin of the rising motion is negative climb due to aggregation of interstitial atoms along with dislocations. This is based on a consideration that interstitial Si or Ge atoms may be generated at the oxidation front beneath the surface oxide layer, and then quickly diffuse into the SGOI layer at high temperature. In other words, it is thought that the rising up of the fragments of MDs is related to the supply of interstitials by the oxidation of the surface, and leads to the incidence of TDs. Therefore control of the oxygen concentration during the initial stage is important for reducing TD density. If the supply of interstitials could be suppressed, the misfit fragments sink toward the BOX interface and would disappear among the interface. The rising up of the misfit fragments mainly takes place during the ramping-up period that is followed by the high-temperature oxidation and condensation process. Therefore, the suppression of oxidation during the ramping-up period of the oxidation reactor is required for formation of the SGOI with few dislocations; however, the slow oxidation is mandatory for the protection of the fragile SiGe surface under the high-temperature condition.

In order to solve this problem, the two-step oxidation and condensation method, which combines the pre-oxidation to form a thin capping oxide layer for protecting the SiGe surface, with a temperature ramping period under no oxygen atmosphere, is proposed [34]. Experimentally, SiGe layers with a Ge content of 10% on SOI substrates are prepared as starting substrates for the Ge condensation process. Here, the initial thickness of the SiGe epitaxial layer is 250 or 400 nm. At first, the 15 nm oxide layer is formed at 900°C and the temperature rises to 1200°C without oxygen. It has already been confirmed that the 15 nm oxide can protect SiGe surfaces from the degradation during ramping up toward 1200°C. After that, the Ge condensation process proceeds at 1200°C. Figure 8.9 shows a comparison of plan-view and cross-sectional TEM images of SGOI layers formed by (a) the conventional Ge condensation method and (b) the two-step method. It is clearly observed that the rising up of misfit fragments is suppressed by the two-step method. The TD density of the SGOI layers formed by the two-step method is quantitatively evaluated by the enhanced secco etching technique [27, 28]. As a result, an etch pit density (EPD) value as low as 1×10^3 cm^{-2} is obtained at an etching depth of less than 100 nm from the surface.

An EPD value of 2×10^3 cm^{-2} has also been obtained in the case of a SGOI sample having 230 nm thickness and Ge content of 20% with the full


(a) Conventional method
(b) Two-step method
1 µm
1 µm
Misfit fragments have not risen up
300 nm
500 nm
5 min. after the temperature reached 1200°C

8.9 Comparison of plan-view and cross-sectional TEM images of SGOI layers formed by (a) the conventional Ge condensation method and (b) the two-step method.

relaxation; however, the EPD value of the SGOI sample having a similar structure formed by the conventional condensation method is 1×10^8 cm^{-2}. It has also been confirmed that the EPD value is 3×10^7 cm^{-2} when the sample having a capping oxide layer has been exposed to an oxygen atmosphere during the ramping-up period [31]. These results prove that ramping up under no oxygen is effective for reducing the TD density of SGOI formation.

8.2.4 Variation of fabrication method of SGOI and strained-Si on insulator (sSOI)

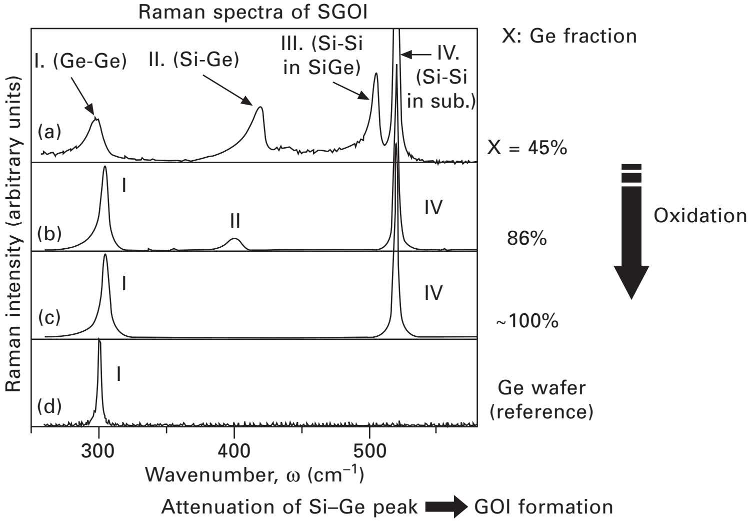
In the previous sections, the development of the Ge condensation method, which has been led by the simple principles of rejection of Ge atoms from the oxide layer during oxidation of silicon and the diffusion barrier effect for Ge atoms by the oxide layer, is described as a technology for fabricating the SGOI structure. Next, the other method of forming the SGOI structure is introduced. Because the SGOI structure is recognizably a variation of the SOI structure, it is possible to apply a SOI fabrication method that is similar to the SGOI fabrication method. In other words, the bonding method, which is the conventional method for SOI fabrication, can be applied for SGOI fabrication [35]. In the bonding technology for SGOI fabrication, the thick and relaxed SiGe epitaxial layer grown on a bulk Si substrate is prepared, and the surface of the SiGe epitaxial layer is flattened by the chemical–mechanical polishing (CMP) method. The SiGe surface is bonded on the surface oxide layer of the base substrate, and then finally the SGOI layer is completed after the back-substrate of the SiGe epitaxial layer is etched off. Here, the stripping method known as smart cut, which consists of high-dose hydrogen ion implantation and cutting at the implanted portion, is also applicable to the cutting of the SiGe layer. For the bonding method, the limitation of the Ge content that applies in the Ge condensation method does not exist.

As mentioned in the introduction, one of the purposes of SGOI fabrication is to obtain the template of the strained Si layer. However, it is possible to obtain the strained Si layer directly laminated on the buried oxide layer by the bonding method [36]. Specifically, the strained Si layer is grown on the SiGe epitaxial layer, which has been grown on the bulk Si substrate, and the strained Si layer is bonded on the Si oxide layer on the base substrate. The entire SiGe layer including the Si substrate is removed, leaving only the strained Si layer, and finally the strained-Si on insulator (sSOI) structure, which is very similar to the conventional SOI structure except for the strain in the layer, is completed. It should be noted that the strain remains even after the stripping of the SiGe layer, which has worked as the stressor, because the bonding between the strained Si layer and the oxide layer is so tight. However, it is also reported that the strain in the Si layer is partially relaxed after the high-temperature thermal process and/or the fine patterning process [37].

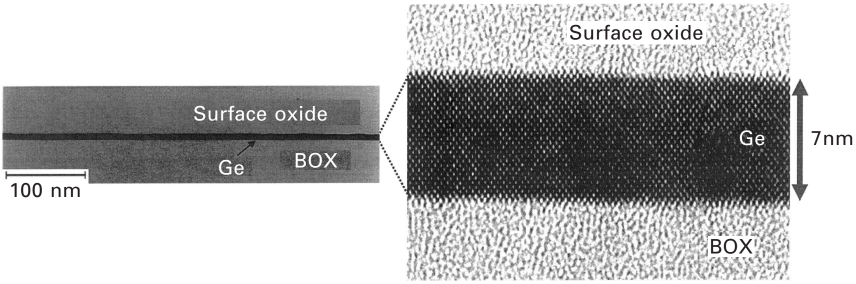
8.3 Extension toward Ge on insulator

Continuation of Ge condensation can simply lead to a pure germanium on insulator (GOI) structure as a result of oxidizing all the Si atoms included in the SiGe layers. The formation of this high-purity GOI layer is attributable to the combination of the selective oxidation of Si atoms in the SGOI layer and the fast diffusion of Si atoms in SiGe crystal with high Ge contents. If the oxidation condition is well tuned, the supply of silicon atoms to the oxide interface can be kept larger than the consumption of silicon atoms by oxidation, and the silicon atoms continue to be consumed until all atoms are oxidized. It has already been reported that GOI layers with an ultra-thin body as thin as 7 nm can be formed [38] by the Ge condensation method and the fabrication technology is summarized in the following. Experimentally, fabrication of the GOI structure has been achieved by the same method as in the case of Ge condensation for the SGOI structure, that is, the dry oxidation of the SiGe epitaxial layer grown on the SOI substrate. During this oxidation process, the melting point of SiGe decreases with the increase of the Ge content, and the oxidation temperature has to be kept below the melting point by reducing the furnace temperature. As mentioned in the previous section, there is a risk of the formation of an abrupt Ge profile at the SGOI/BOX interface, because the decrease in oxidation temperature brings about a decrease of the diffusion rate of Ge atoms in SiGe crystal, which is more effective than a decrease of the oxidation rate. However, it is possible to proceed with the Ge condensation process toward GOI formation, because the diffusion coefficient of Ge atoms in SiGe crystal is greatly enhanced with the increase of Ge content in SiGe [39].

One of the most important issues in the Ge condensation process for GOI formation is the lack of an auto-stop mechanism for the oxidation. Ge atoms are suddenly oxidized when all the silicon atoms are consumed and the oxide quality deteriorates. Therefore, for obtaining excellent GOI, it is important to stop oxidation just as consumption of Si atoms ends. In order to define the end point, analysis of Raman spectroscopy is useful. Figure 8.10 shows the variation of the Raman spectroscopy during the Ge condensation process. In the figure, the peaks denoted I, II, III, and IV are attributed to the vibration mode of the Ge–Ge bond, the vibration mode of the Si–Ge bond, the vibration mode of the Si–Si bond in SiGe crystal, and the vibration mode of the Si–Si bond in Si substrate, respectively. Peaks II and III, which are related to the Si atoms in the SiGe crystal, decrease as the Ge condensation proceeds and finally disappear. Figure 8.11 shows cross-sectional TEM images of a GOI layer 7 μm thick. The lattice image of the Ge crystal is clearly observed



8.10 Variation of the Raman spectroscopy during the Ge condensation process.



8.11 Cross-sectional TEM images of GOI layer 7 nm thick.

and also reveals that the interface between the Ge layer and the oxide layer is flat. It has been determined that the residual Si concentration in the GOI layer is less than 0.01% by SIMS analysis.

8.4 Conclusion

The formation technology of the SGOI structure has been explained. Starting from the application of the SIMOX technique on the SiGe layer, an advanced method called the Ge condensation method has been developed to obtain the SGOI for the template substrate for the strained Si layer. Various physical properties such as relaxation and control of strain, the formation mechanism of dislocations and the reduction of dislocations have also been discussed with a view to improving the Ge condensation technology. It is revealed that not only the SGOI structure but also the Ge on insulator (GOI) structure can be obtained by the extension of the Ge condensation method. Application of the Ge condensation method is expected to extend further in the future.

8.5 Acknowledgment

Research and development regarding Ge condensation and the SIMOX process for SGOI formation were supported by NEDO and the MIRAI project.

8.6 References

[1] G. Abstreiter, H. Brugger, T. Wolf, H. Jorke and H.J. Herzog (1985) *Phys. Rev. Lett.* **54**, 2441.

[2] D.K. Nayak, J.C.S. Woo, J.S. Park, K.L. Wang and K.P. MacWilliams (1993) *Appl. Phys. Lett.* **62**, 2853.

[3] H. Daembkes, H.J. Herzog, H. Jorke, H. Kibbel and E. Kasper (1986) *IEEE Trans. Electron Devices* **ED-33**, 633.

[4] E.A. Fitzgerald, Y.-H. Xie, D. Monroe, P.J. Silverman, J.M. Kuo, A.R. Kortan, F.A. Thiel and B.E. Weir (1992) *J. Vac. Sci. Technol.* **B10**, 1807.

[5] D. Dutartre, P. Warren, F. Provenier, F. Chollet and A. Perio (1994) *J. Vac. Sci. Technol.* **A12**, 1009.

[6] P.M. Mooney, J.L. Jordan-Sweet, K. Ismail, J.O. Chu, R.M. Feenstra and F.K. LeGoues (1995) *Appl. Phys. Lett.* **67**, 2373.

[7] J.H. Li, G. Springholz, J. Stangl, H. Seyringer, V. Holy, F. Schaeffler and G. Bauer (1998) *J. Vac. Sci. Technol.* **B16**, 1610.

[8] A.C. Churchill, D.J. Robbins, D.J. Wallis, N. Griffin, D.J. Paul, A.J. Pidduck, W.Y. Leong and G.M. Willams (1998) *J. Vac. Sci. Technol.* **B16**, 1634.

[9] J.L. Liu, C.D. Moore, G.D. U'Ren, Y.H. Luo, Y. Lu, G. Jin, S.G. Thomas, M.S. Goorsky and K.L. Wang (1999) *Appl. Phys. Lett.* **75**, 1586.

[10] A.R. Powell, S.S. Iyer and F.K. LeGoues (1994) *Appl. Phys. Lett.* **64**, 1856.

[11] N. Sugiyama, T. Mizuno, S. Takagi, M. Koike and A. Kurobe (2000) *Thin Solid Films* **369**, 199.

[12] T. Mizuno, S. Takagi, N. Sugiyama, H. Satake, A. Kurobe and A. Toriumi (2000) *IEEE Electron Device Lett.* **EDL-21**, 230.

[13] N. Sugiyama, T. Mizuno, M. Suzuki and S. Takagi (2001) *Jpn. J. Appl. Phys.* **40**, 2875.

[14] T. Tezuka, N. Sugiyama, T. Mizuno, M. Suzuki and S. Takagi (2001) *Jpn. J. Appl. Phys.* **40**, 2866.

[15] N. Sugiyama, T. Tezuka, T. Mizuno, M. Suzuki, Y. Ishikawa, N. Shibata and S. Takagi (2004) *J. Appl. Phys.* **95**, 4007.

[16] R.W. Olesinski and G.J. Abbascian (1984) *Bull. Alloy Phase Diagrams* **5**, 180.

[17] F.K. LeGoues, R. Rosenberg, T. Nguyen, F. Himpsel and B.S. Meyerson (1989) *J. Appl. Phys.* **65**, 1724.

[18] F.J. Norton (1961) *Nature* **191**, 701.

[19] M. Ogino, Y. Onabe and M. Watanabe (1982) *Phys. Stat. Sol. (a)* **72**, 535.

[20] T. Tezuka, N. Sugiyama and S. Takagi (2001) *Appl. Phys. Lett.*, **79**, 1798.

[21] J.C. Tsang, P.M. Mooney, F. Dacol and J.O. Chu (1994) *J. Appl. Phys.*, **75**, 8098.

[22] S. Nakashima, T. Mitani, M. Ninomiya and K. Matsumoto (2006) *J. Appl. Phys.* **99**, 053512.
[23] T. Tezuka, N. Sugiyama, S. Takagi and T. Kawakubo (2002) *Appl. Phys. Lett.*, **80** 3560.
[24] T. Tezuka, N. Sugiyama and S. Takagi (2003) *J. Appl. Phys.* **94**, 7553.
[25] A. Nishida, K. Nakagawa, E. Murakami and M. Miyao (1992) *J. Appl. Phys.* **71**, 5913.
[26] H. Yin, R. Huang, K.D. Hobart, Z. Suo, T.S. Kuan, C.K. Inoki, S.R. Shieh, T.S. Duffy, F.J. Kub and J.C. Sturm (2002) *J. Appl. Phys.* **91**, 9716.
[27] J. Margail, J.M. Lamure and A.M. Papon (1992) *Mater. Sci. Eng.* **B12**, 27.
[28] N. Sugiyama, Y. Moriyama, T. Tezuka, T. Mizuno, S. Nakaharai, K. Usuda and S. Takagi (2003) *Jpn. J. Appl. Phys.* **42**, 4476.
[29] K.W. Schwarz (1999) *J. Appl. Phys.* **85**, 120.
[30] Z. Yang, J. Alperin, W.I. Wang and S.S. Iyer (1998) *J. Vac. Sci. Technol.*, **B16**, 1489.
[31] E.M. Rehder, C.K. Inoki, T.S. Kuan and T.F. Kuech (2003) *J. Appl. Phys.* **94**, 7892.
[32] Y.H. Lo (1991) *Appl. Phys. Lett.* **59**, 2311.
[33] F.K. LeGoues, A. Powell, S.S. Iyer (1994) *J. Appl. Phys.* **75**, 7240.
[34] N. Sugiyama, S. Nakaharai, N. Hirashita, T. Tezuka, Y. Moriyama, K. Usuda and S. Takagi (2007) *Semicond. Sci. Technol.* **22**, S59.
[35] G. Taraschi, T.A. Langdo, M.T. Currie, E. Fitzgerald and D.A. Antniadis (2002) *J. Vac. Sci. Technol.* **B20**, 725.
[36] T.A. Langdo, M.T. Currie, A. Lochtefeld, R. Hammond, J.A. Carlin, M. Erdtmann, B. Braithwaite, V.K. Yang, H. Badawi and M.T. Bulsara (2003) *Appl. Phys. Lett.*, **82**, 4256.
[37] M. Erdtmann, T.A. Langdo, C.J. Vineis, H. Badawi and M.T. Bulsara (2003) *Extended Abstracts of 2003 International Conference on Solid State Device and Material*, Tokyo, p. 290.
[38] S. Nakaharai, T. Tezuka, N. Sugiyama, Y. Moriyama and S. Takagi (2003) *Appl. Phys. Lett*, **83**, 3516.
[39] S. Nakaharai, T. Tezuka, N. Hirashita, E. Toyoda, Y. Moriyama, N. Sugiyama and S. Takagi (2009), *J. Appl. Phys.* **105**, 024515.

9

Miscellaneous methods and materials for silicon–germanium (SiGe) based heterostructures

M. MIYAO, Kyushu University, Japan

Abstract: Si-based heterostructure technologies have been widely developed over the last quarter century to overcome the scaling limit of CMOS performance. Strained growth, modulation doping, and formation of quantum well structures developed on Si substrates are the typical examples, which achieved significant enhancement of carrier mobility and electroluminescence efficiency. What is the next jump? Advanced crystal growth of high-quality SiGe crystals on insulating films is expected to open up high-performance system-in-displays and three-dimensional stacked devices. New functions created by spin-injection from ferromagnetic electrodes into SiGe channels are other candidates to be used for this purpose. In line with this, the present chapter describes recent developments in miscellaneous crystal growth methods and materials essential for SiGe on insulator structures and SiGe spintronics.

Key words: three-dimensional stacked devices, thin-film transistors, SiGe spintronics, SiGe on insulaor, ferromagnetic silicide, high carrier mobility, spin injection, metal-induced lateral crystallization, electric-field assisted metal-induced lateral crystallization, seeding rapid-melting-growth, SiGe-mixing-triggered rapid melting growth, metal-imprint techniques

9.1 Introduction

Research and development for new semiconductor devices which enable ultra-high speed operation, ultra-low power dissipation, and/or multi-functional operation are strongly required to overcome the scaling limit of complementary metal-oxide-semiconductor (CMOS) performance. In line with this, Si-based heterostructure technologies have been widely developed in the past quarter-century. Among them, development of the SiGe heteroepitaxy technique on Si substrates has enabled strained growth, modulation doping, and formation of quantum well structures, which achieved significant enhancement of carrier mobility and electroluminescence efficiency [1–5]. And today, strained Si epitaxially grown on relaxed SiGe buffer layers becomes the mainstream of CMOS structures [6–8].

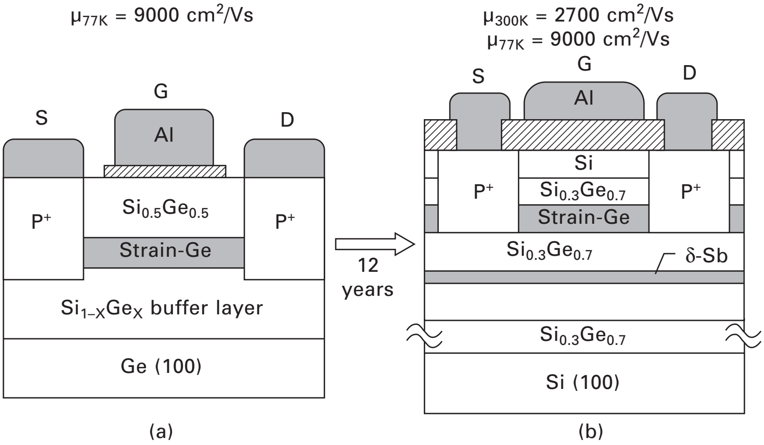
What is the next jump? Advanced crystal growth of high-quality Si and SiGe crystals on insulating films is expected to open up high-performance system-in-displays and three-dimensional (3-D) stacked devices [9–13]. New functions created by spin injection from ferromagnetic electrodes into SiGe channels or quantum dots are other candidates to be used for this purpose [14–17].

This chapter describes recent developments in miscellaneous crystal growth methods and materials essential for SiGe on insulator (SGOI) structures and SiGe spintronics.

9.2 Oriented growth of silicon-germanium (SiGe) on insulating films for thin film transistors and 3-D stacked devices

Figure 9.1(a) shows the schematic device structure reported by the author in 1991 [2]. Very high hole mobility of 9000 $cm^2/(Vs)$ was achieved at 77 K by developing a strained Ge channel. However, the device structure was limited on the Ge substrate. After 12 years, Irisawa *et al*. fabricated the strained Ge channel on the Si substrate [4], as shown in Fig. 9.1(b). This means that it took as long as 12 years to develop the SiGe virtual substrate technology in order to implement the device on the Si substrate, and moreover, the strained Ge channel technology is very attractive and worth such a huge development cost. The technological difficulty due to the difference in the lattice constants between Si (0.117 nm) and Ge (0.122 nm) is eliminated for thin-film transistors (TFTs) fabricated on glass or quartz substrates. Consequently, the Ge or SiGe channel should be used for future TFT.

Recently, Schottky source/drain (S/D) MOS field-effect transistors (FETs) fabricated on poly-Ge/SiO_2 structure were reported [18]. The kink effect due



(a) (b)

9.1 Progress of ultra-high-speed Ge channel devices: (a) Miyao *et al*., 1991 [2]; (b) Irisawa *et al*., 2003 [4].

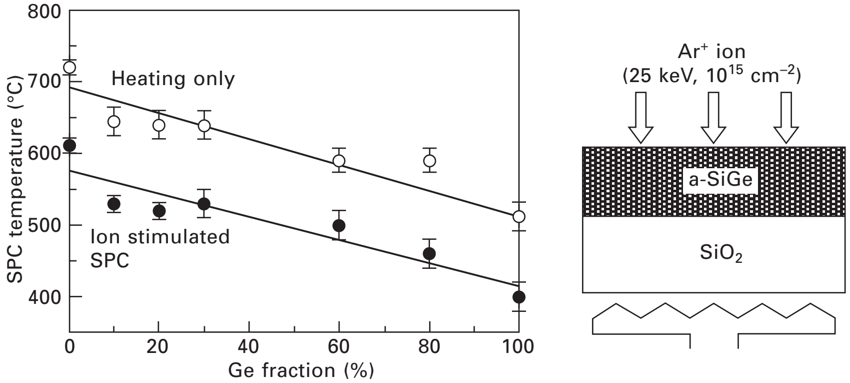
to the floating body effect, which was observed in the conventional doping S/D TFTs, was successfully suppressed. However, hole mobility was limited to 100 cm^2/Vs. In order to enhance the carrier mobility, new crystal growth methods enabling oriented growth of SiGe on insulating films are strongly required.

In line with this, the present section reports on recent progress in solid-phase crystallization (SPC) [19–22], metal-induced lateral crystallization (MILC) [23–27], and liquid-phase crystallization (LPC) of amorphous-SiGe (a-SiGe) [13, 28–34] on insulating films.

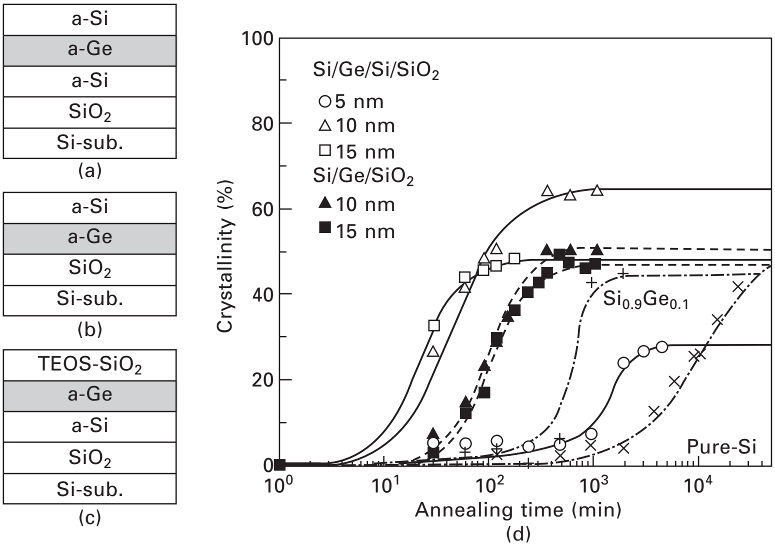
9.2.1 Enhanced crystal nucleation for low-temperature solid-phase crystallization

SPC of a-SiGe is initiated by crystal nucleation and followed by nuclei growth. Consequently, nucleation control is the key factor to control SPC. The SPC temperature of SiGe, i.e., the critical temperature for crystal nucleation, is summarized in Fig. 9.2 as a function of the Ge fraction [19, 20]. Here, a-$Si_{1-x}Ge_x$ films (x = 0–1, thickness 90 nm) deposited on SiO_2 substrates were annealed (20 min) in a vacuum. The SPC temperature clearly decreases with increasing Ge fraction. Ge doping into Si reduces the bond strength between Si and Si, i.e., Si–Si 3.73 eV, Si–Ge 3.65 eV and Ge–Ge 3.56 eV, which enhances the bond rearrangement, i.e. crystal nucleation. These results suggest that local Ge insertion into a-Si films should be effective to enhance SPC of a-Si, where crystal nucleation is initiated in local Ge layers and then nuclei growth propagates into a-Si layers.

Three types of stacked structures [21], i.e., (a) a-Si/a-Ge/a-Si/SiO_2, (b) a-Si/a-Ge/SiO_2, and (c) SiO_2/a-Ge/a-Si/SiO_2, are shown in Figs 9.3(a), (b), and



9.2 The SPC temperature of a-SiGe on SiO_2 with and without ion irradiation as a function of Ge fraction.



9.3 Schematic cross-section of sample structures: (a) a-Si/a-Ge/a-Si/SiO$_2$, (b) a-Si/a-Ge/SiO$_2$, and (c) SiO$_2$/a-Ge/a-Si/SiO$_2$. (d) Crystallinity change during isothermal annealing (600°C). The results obtained for a-Si/SiO$_2$ and a-Si$_{0.9}$Ge$_{0.1}$/SiO$_2$ are also shown for reference.

(c), respectively, where a-Ge thicknesses are chosen as 5–15 nm. Isothermal annealing characteristics (600°C) of crystallinity are summarized in Fig. 9.3(d). The Ge depth profiles of those samples were measured by Auger electron spectroscopy. In structure (a), Ge atoms diffused into a-Si layers, and no change of SPC was detected. In structures (b) and (c), Ge atoms were localized in the initial position, which caused significant enhancement of SPC. These results indicate that two-dimensional nucleation is initiated in Ge layers directly contacted to SiO$_2$, and then nuclei growth propagates into a-Si layers. In this way, enhanced SPC of a-Si is achieved by fabricating the sandwiched structures of (b) and (c).

Ion beam irradiation [20], micro-indentation [35], and metal-imprint [36] techniques are effective in enhancing crystal nucleation in a-SiGe with all Ge fractions (0–100%). The SPC temperature of a-SiGe under Ar$^+$ ion-beam irradiation (25 keV, ~10^{15} cm^{-2}) is summarized in Fig. 9.2. The critical temperature for SPC is successfully decreased by Ar$^+$ irradiation, e.g. from 650°C (without irradiation) to 450°C (with irradiation) for a-Si$_{0.7}$Ge$_{0.3}$. This temperature is also decreased with increasing Ge fraction. Consequently, the SPC temperature can be decreased by 150°C for a-SiGe with all Ge fractions, and growth below the softening temperature of soda-lime glass substrates (450°C) is achieved. Generation of point defects by Ar$^+$ irradiation is considered to decrease the energy necessary for the rearrangement of

Si–Si, Si–Ge and Ge–Ge bonds, which enhances the crystal nucleation at a low temperature (~400°C).

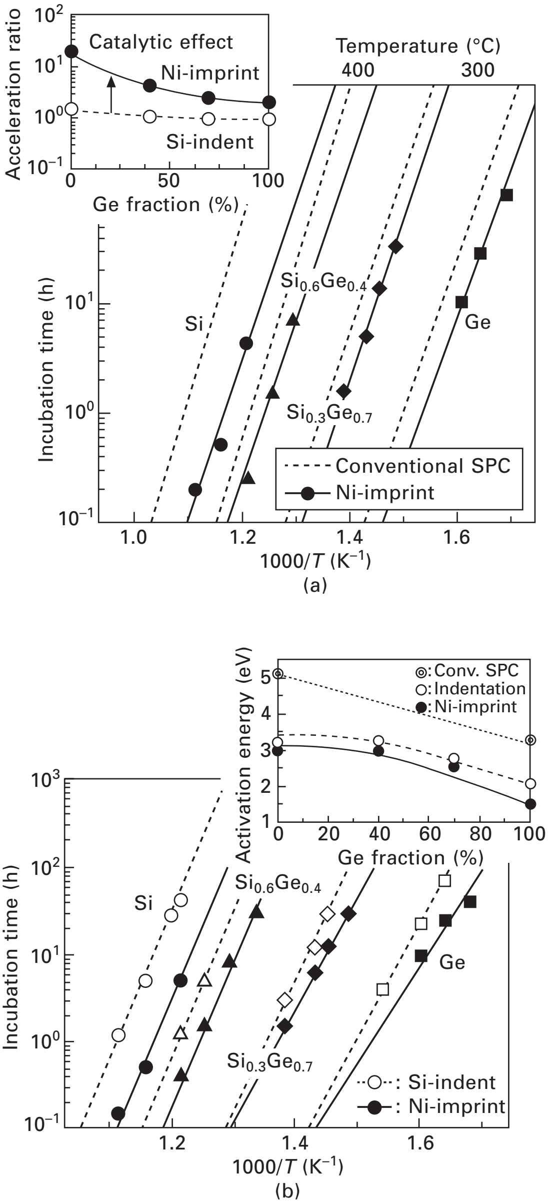
Arrhenius plots of the incubation time of a-SiGe crystal growth [35, 36] are summarized in Figs 9.4(a) and (b), where the Ni-imprint assisted SPC process and micro-indentation assisted SPC are compared with the conventional SPC process. The incubation times are significantly reduced by adding these external processes to the conventional heating-only process. Enhanced nucleations are attributed to the catalytic effect induced by metal-imprint and the stress effect induced by indentation, respectively. Lateral SPC growth propagates preferentially around these crystal nuclei. As a result, SiGe crystal arrays with diameters of a few microns are formed at the imprinted or indented position.

9.2.2 Metal-induced lateral crystallization (MILC) for SiGe with large grain

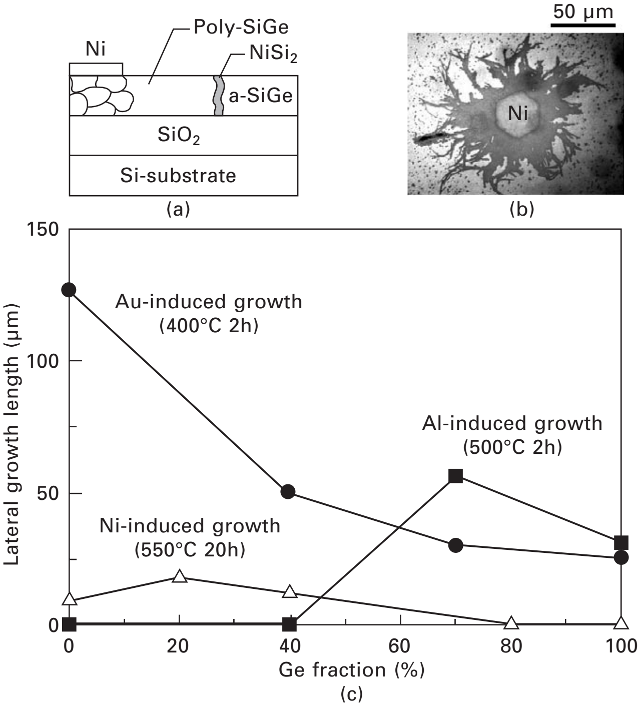
Metal-induced lateral crystallization (MILC) is a low-temperature crystal growth using the catalytic effect of some metals, i.e., Al, Pd, Ag, Cu, and Ni [12, 23, 26]. TFT operation with high mobility has already been demonstrated using the high-quality poly-Si films formed by MILC. In order to open new windows for advanced TFTs with the SiGe channel, the MILC process in SiGe has been investigated [23, 24].

Patterned Ni films (15 nm thick) are evaporated on a-$Si_{1-x}Ge_x$ layers ($x = 0$–1, 50 nm thickness) deposited on SiO_2/Si structures [23], as shown in Fig. 9.5(a). Figure 9.5(b) shows a typical Nomarski optical micrograph after annealing (550°C, 20 h), which indicates $Si_{0.4}Ge_{0.6}$ lateral growth around the hexagonal Ni pattern. It was found that crystallized features strongly depended on the Ge fraction. In the case of low Ge fractions ($x = 0$–0.2), the crystallized regions showed plane morphology. This suggests that the mechanism of MILC in a-SiGe ($x = 0$–0.2) is very close to that in a-Si. When the Ge fraction exceeded 0.4, the features of MILC were drastically changed. Dendrite crystallization became dominant. Originating from the difference in atomic radii between Si and Ge, a local strain should be induced during MILC. Consequently, Ni atoms are considered to migrate preferentially along such a strained field, which results in dendrite crystallization.

Effects of other types of catalysis such as Al and Au have been examined to obtain plane growth for a-$Si_{1-x}Ge_x$ with all Ge fractions ($x = 0$–1). This is because these metals form eutectics with a-Si and a-Ge at a low temperature. The plane growth length in Al-induced crystallization (500°C, 2 h) is shown as a function of the Ge fraction in Fig. 9.5(c), together with the data for Ni (550°C, 20 h). The results for Al show that for high Ge fractions ($x = 0.7$–1), lateral growth proceeds faster than that for Ni. However, growth does not proceed for low Ge fractions ($x = 0$–0.4). It is speculated that the Ge fraction



9.4 (a) Arrhenius plots of the incubation times for nucleation in a-SiGe (solid line: Ni-imprinted positions, dashed line: non-imprinted regions). Insert: Ge fraction dependence of the acceleration ratios of nucleation. (b) Arrhenius plots of the incubation times for nucleation in a-SiGe (open symbols: indentation-induced SPC, closed symbols: Ni-imprint induced SPC). Insert: Ge fraction dependence of activation energy for nucleation. Plots for the different proportions of Si and Ge are further distinguished by the use of different shapes of symbols.



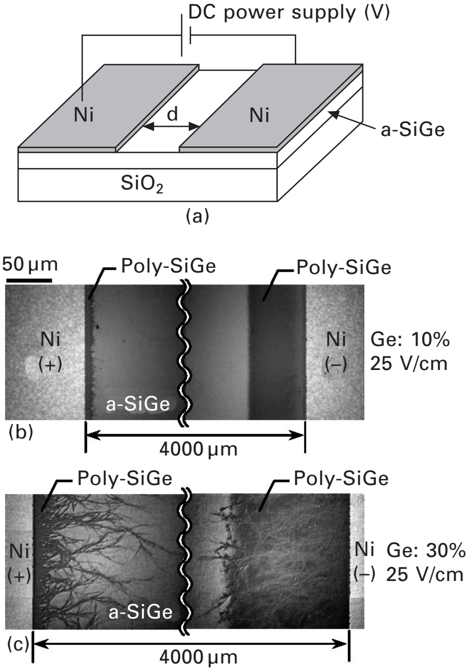
(a) (b) (c)

9.5 (a) Schematic cross-section of sample structure. (b) Nomarski optical micrograph of SiGe samples (Ge 60%) after annealing at 550°C. (c) Ge fraction dependence of lateral growth length for SiGe samples grown with Al, Au, and Ni as catalysts.

dependence is due to the different eutectic temperatures for Al–Si (577°C) and Al–Ge (420°C). Since the eutectic temperatures for Au–Si (363°C) and Au–Ge (361°C) are very similar, the Au-induced crystallization (400°C, 2 h) is examined and summarized in Fig. 9.5(c). Uniform crystallization with long growth length (>30 μm) is obtained for all Ge fractions (x = 0–1) at 400°C. However, a large amount of Au (10–15%) remains in the grown SiGe layers, which is a significant drawback originating from the eutectic reaction.

9.2.3 Electric-field assisted MILC for uniform and nanowire growth

Application of electric fields during Ni-MILC is effective to achieve plane growth of a-SiGe with high Ge fractions [25]. Figure 9.6 shows a sample structure and Nomarski optical micrographs after annealing (525°C, 25 h). The patterned Ni films deposited on a-SiGe were employed as the electrodes

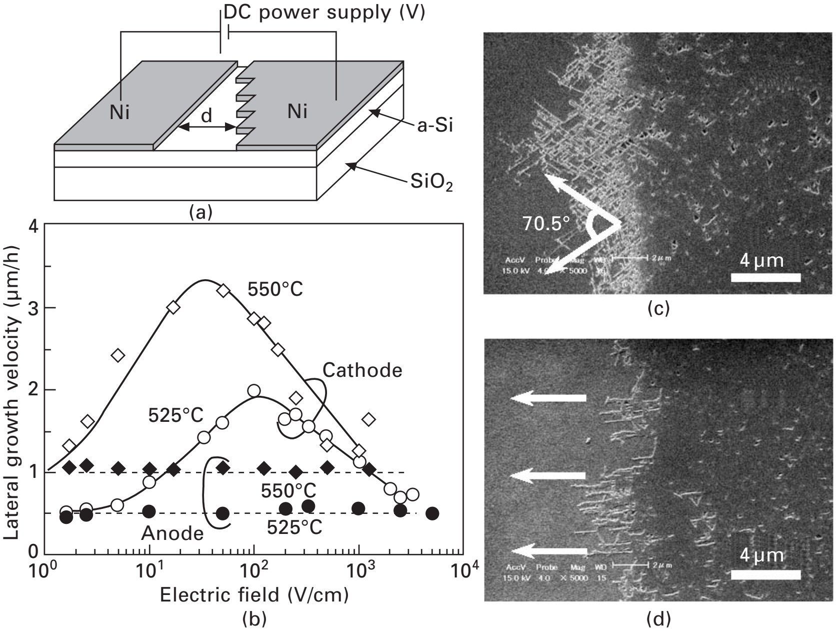


9.6 (a) Schematic sample structure; (b), (c) optical micrographs of SiGe samples after MILC (525°C) under electric fields.

for DC power supply. Although the growth at the anode did not depend on the applied field, the growth at the cathode is significantly enhanced by applying the field. Jang *et al*. reported that Ni atoms in $NiSi_2$ are negatively charged [37], which suggests that migration of charged Ni atoms in SiGe is enhanced by an electric field. In this way, uniform SiGe crystallization becomes possible around Ni patterns even for samples with a high Ge fraction. Growth lengths sufficient for device fabrication (>100 μm) are obtained.

Now, let us discuss the possibility of directional nanowire formation by MILC growth. Since carrier scattering is expected to be reduced by two-dimensional carrier confinement in the nanowire crystal regions, such structures should be used as the channel regions for advanced high-performance TFTs. The dendrite growth characteristics observed in Ni-induced crystallization of a-$Si_{1-x}Ge_x$ ($x = 0.4$–0.6) (see Fig. 9.5(b)) were investigated as a function of the annealing temperature. Observation by scanning electron microscopy (SEM) indicated that the dendrite crystal regions became straight, and the width became narrower with decreasing annealing temperature. The widths of dendrite regions were estimated to be 0.5 μm (550°C), 0.2 μm (500°C), and 0.05 μm (450°C). In addition, it was found that the growth length exceeded 10 μm. Consequently, very sharp needle-like crystal regions were obtained by lowering the annealing temperature. However, the growth direction of the needle-like crystal regions was random. In order to apply the sharp needle-like crystals to devices, the crystal growth direction should be controlled.

Electric field-assisted MILC is very useful for this purpose [25]. The sample structure and the lateral growth velocity (525°C, 550°C) at the cathode obtained under a wide range of electric fields (2–3000 V/cm) are shown in Figs 9.7(a) and (b), respectively. The growth velocity increases with increasing electric field up to 100 V/cm, and shows a peak. A SEM micrograph of the growth regions (525°C, 25 h) around the Ni cathode under the electric field of 33 V/cm is shown in Fig. 9.7(c). Dendrites with random directions are observed around the growth front. Over the electric field of 100 V/cm, the growth velocity decreases with increasing electric field, as shown in Fig. 9.7(b). The decrease in the growth velocity at the higher electric field (>100 V/cm) may be due to the depletion of Ni^- ions at the growth front. Figure 9.7(d) shows a SEM micrograph of the crystallized regions (525°C, 25 h) under the electric field of 2500 V/cm. It is found that the growth direction of dendrites aligns to the high electric field. This is a big advance for controlling the growth direction of nanowire channel regions for fabrication of future TFTs.



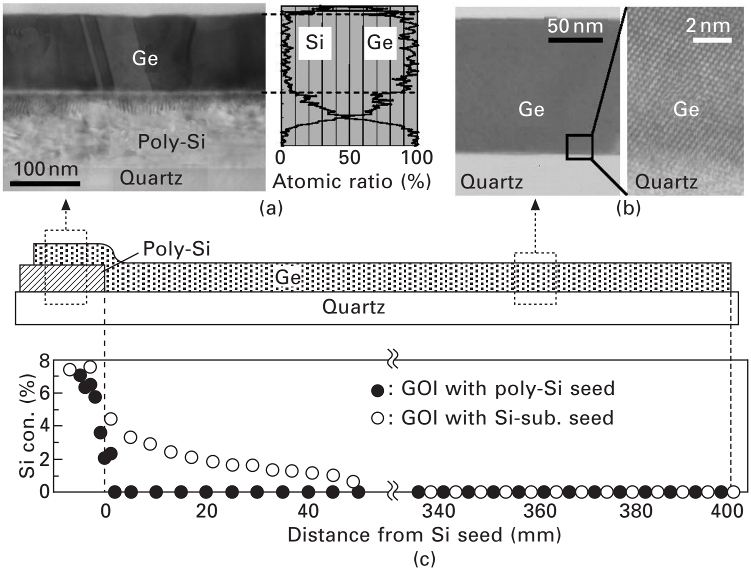
9.7 (a) Schematics of experimental setup; (b) electric field dependence of lateral growth velocity for Si samples; (c), (d) micrographs of growth front at the cathode side after MILC (525°C) under electric fields of (c) 33 V/cm, and (d) 2500 V/cm.

9.2.4 SiGe-mixing-triggered rapid melting growth for defect-free Ge on insulator (GOI)

Seeding of the rapid melting growth of a-Ge [29, 32], similar to the laser-induced lateral growth of Si [38], has been developed to obtain single-crystal Ge on insulator (GOI) structures, where a-Ge layers deposited on SiO_2 films were first grown vertically from Si substrates through opening windows formed in SiO_2 films, and then propagated laterally over SiO_2. These efforts achieved defect-free single-crystal Ge wires (20–40 μm in length, 2–3 μm in width) on insulating films. However, this method has not been applied to TFT fabrication, because Si substrates were considered to be essential to initiate the seeding of rapid melting growth.

Recently, the details of this process were clarified, in that the driving force to cause the lateral growth of Ge was not the thermal flow from the molten Ge to the Si substrates through seeding windows, but the spatial gradient of the solidification temperature originating from Si–Ge mixing in seeding areas [29]. This triggered a new idea of the directional lateral growth of a-Ge on transparent insulating substrates [13].

A typical example of rapid melting growth of a-Ge on quartz substrates using polycrystalline Si (poly-Si) islands as the seed is shown in Fig. 9.8,



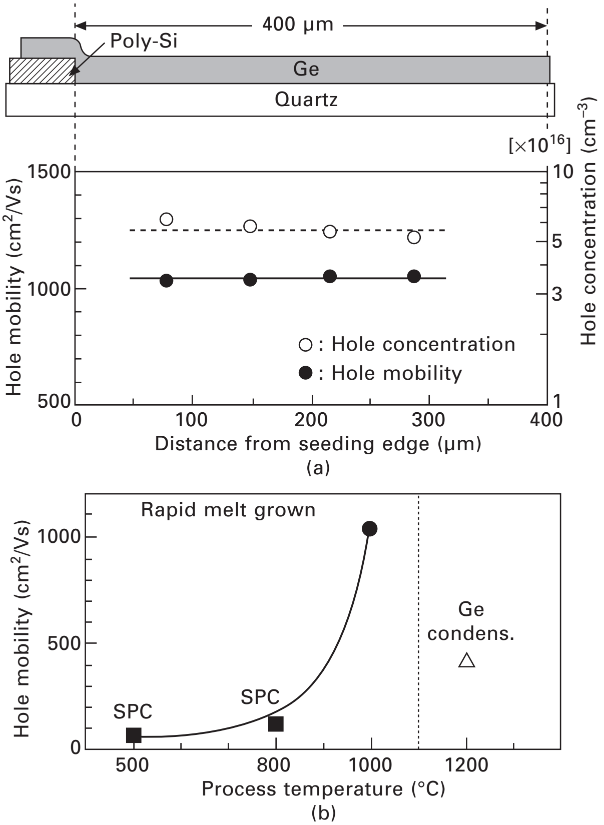
9.8 (a) X-TEM image and depth distribution of Si and Ge concentration in the seeding region; (b) X-TEM image in the grown region over 350 μm away from the seeding area; (c) Si concentration obtained from Raman spectra as a function of the distance from the seeding edge.

where crystal structures of the grown layers and depth distributions of Si and Ge concentrations in the seeding area are displayed. These results indicate the Si–Ge mixing and high defect density at the Si/Ge interface region in the seeding area, and single-crystal Ge without dislocation or stacking fault in the lateral-growth region. Dislocations originating in the interface region of a-Ge and poly-Si are considered to escape to the sample surface along the plane normal to the contour of the grading solidification temperature. Such a defect-necking process has been shown in the previously reported rapid melting growth experiments using a-Ge/SiO_2 structures with bulk-Si seeds (Si substrates) [29, 32, 33].

Electron backscattering diffraction (EBSD) measurements for many samples indicated that the dominant crystal orientation (>50%) of laterally grown Ge was (100), which agreed well with that of crystal grains in the poly-Si islands used as seeds. This demonstrates that the growth is initiated from the poly-Si islands and propagates laterally over 400 μm length. The lateral distribution of the Si concentration is summarized in Fig. 9.8(c), where results obtained from the samples with the bulk-Si seed (Si substrate) are also shown [29]. The Si concentration in the sample on a quartz substrate decreases abruptly along the growth direction compared to that shown in the sample with the bulk-Si seed (open circles). This is one advantage of using thin poly-Si seeds instead of bulk-Si seeds (Si substrates).

The hole concentration and hole mobility measured as a function of the distance from the seeding edge are summarized in Fig. 9.9(a), which indicates uniform electrical properties in the whole grown regions (~400 μm in length). The values for hole mobility in Ge films obtained by various methods are compared in Fig. 9.9(b). The hole mobility of single-crystal Ge obtained by the melt-growth method is significantly higher than that of poly-Ge grown by the SPC process (500°C and 800°C), and even higher than for single-crystal Ge obtained by the oxidation-induced Ge condensation process [8]. In this way, high-quality single-crystal Ge stripes (~400 μm in length) with high carrier mobility (1040 cm^2/Vs) are achieved on transparent insulating substrates of quartz.

To control the crystal orientation of GOI layers formed by rapid melting growth, formation of Si micro-seeds with controlled crystal orientation becomes essential. The metal-imprint assisted SPC process and the Al-induced crystallization (AIC) process are the most useful techniques for this purpose. Very recently, single-crystal (111)-oriented Ge island arrays have been achieved on quartz substrate by using Ni-imprint induced Si(111) micro-crystals as seed for rapid melting growth [30]. In addition, single-crystal (111), (100)-oriented Ge stripes become possible by combining the interfacial-oxide controlled AIC process [31]. These advanced crystal growth methods are now opening new routes to Ge-based TFTs with high-speed operation.



9.9 (a) Hole mobility of melt-grown Ge with poly-Si seed as a function of distance from seeding edge; (b) hole mobility obtained by melt-grown (●), SPC (■), and oxidation-induced Ge condensation (△) process.

9.3 Heteroepitaxial growth of ferromagnetic Heusler alloys for silicon-germanium (SiGe)-based spintronic devices

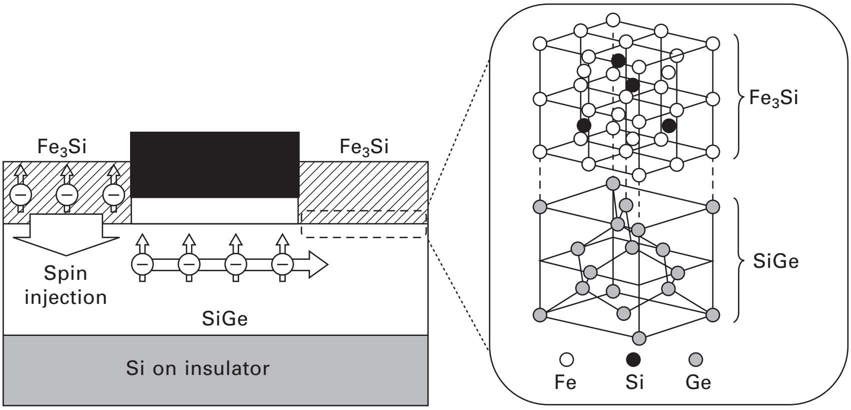
Recently, new approaches to introduce novel functions of light-emitting and spin-transport into CMOS technologies have been explored to overcome the scaling limit of LSI performance [5, 15, 17, 39–41]. To achieve this, heteroepitaxy of miscellaneous materials on SiGe platforms have been widely developed. Fe-based silicide is one of the hot topics in this field. In the last 10 years, efforts have been focused on developing the heteroepitaxy of β-Fe$_2$Si [39], which has achieved a high-efficiency light-emitting diode on the Si platform.

Now interest is expanding to Fe-rich silicide to create spin-transistors by using their ferromagnetic properties. Figure 9.10 shows the typical structure of proposed SiGe-based spin-transistors [15, 17], where the source/drain electrodes consist of ferromagnetic Heusler alloys with high spin polarization and the electrode/SiGe hetero-interfaces become Schottky barriers. DO_3-type ferromagnetic silicide Fe_3Si (Curie temperature 840 K) is calculated to be spin-polarized at the Fermi level [42, 43]. In addition, the lattice constant (0.565 nm) of Fe_3Si is much closed to that of Ge and Si. Consequently, atomically flat interfaces necessary for spin injection are expected in these heterostructures. Another type of spin device becomes possible by developing group IV semiconductor-based dilute magnetic semiconductor (DMS) systems. Chapter 22 in this book describes recent progress in crystal growth and the characteristics of Ge-based DMS (MnGe) for this purpose.

Against such a background, the present section reports on recent progress in the novel epitaxial growth of Heusler alloys on SiGe [44–48]. Magnetic and electrical properties of DO_3-type Fe_3Si grown on SiGe are mainly described. Electrical injection and detection of spin-polarized electrons through the Fe_3Si/Si interface is also demonstrated. In addition, heteroepitaxy of half-metallic alloys ($Fe_{3-x}Mn_xSi$ ($x = 0.6$–1.4), CO_2FeSi) on SiGe substrates is discussed [49–51].

9.3.1 Atomically controlled epitaxial growth of Fe_3Si on Ge substrates

In the heteroepitaxial growth of Fe_3Si, Fe and Si are co-evaporated on the Ge substrates (substrate temperature 60–400°C) using solid-source molecular beam epitaxy (MBE). In-situ reflective high-energy electron diffraction (RHEED) observation indicated that sub-layer (1/3–1/2 ML) coverage with Fe and Si atoms formed amorphous layers covering substrates. This suddenly changed

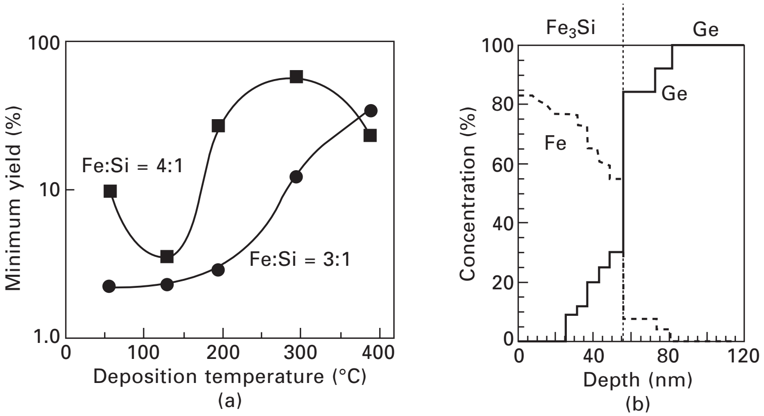


9.10 Schematic structure of spin transistor with SiGe channel.

into the crystalline phase when Fe and Si atoms to form 1 ML Fe_3Si arrived. These phenomena suggest that both Fe/Si ratio and growth temperature are the key factors in obtaining high-quality Fe_3Si [44, 45].

The χ_{min} values of Fe_3Si/(111)Ge structures evaluated from Rutherford backscattering spectroscopy (RBS) are summarized in Fig. 9.11(a) as a function of the Fe/Si ratio and the growth temperature. For off-stoichiometric samples (Fe:Si = 4:1), χ_{min} decreases from 10% to 4% with increasing growth temperature from 60°C to 130°C, where all RHEED patterns showed symmetrical structures of Fe_3Si. When the growth temperature exceeds 130°C, χ_{min} increases up to 60% (300°C), resulting in diffusive RHEED patterns. However, χ_{min} finally decreases to 22% at 400°C, showing a symmetrical structure different from that of Fe_3Si. This suggests that another phase is formed in such a high-temperature region. The depth profiles of Fe and Ge were evaluated by RBS measurements. The hetero-interface was very sharp at 130°C growth; however, atomic diffusion started to occur at 300°C, as shown in Fig. 9.11(b). Finally, the FeSiGe compound layer was formed at 400°C due to the supply of Ge atoms from the substrate. These results suggest that Fe atom diffusion into Ge substrates causes lattice disordering, and consequently its prevention is most important for realizing high-quality interfaces.

The χ_{min} values obtained from stoichiometric samples (Fe:Si = 3:1) shown in Fig. 9.11(a) indicate that a very low χ_{min} value of 2% is achieved for Fe_3Si grown at 130°C. In addition, the temperature dependence of χ_{min} is not so remarkable in the wide temperature range (60–200°C), even though a gradual increase is observed above 300°C. These results suggest that growth at a low temperature (<200°C) under the stoichiometric condition is essential to



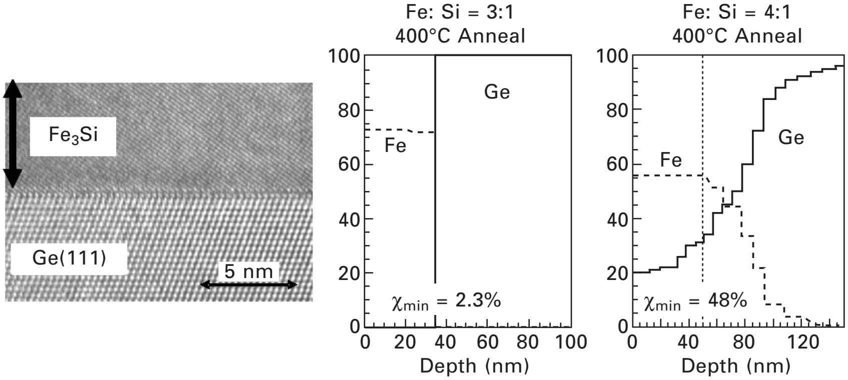
9.11 (a) χ_{min} of Fe_3Si evaluated from RBS spectra as a function of Fe/Si ratio and growth temperature; (b) concentration profiles of Fe and Ge of off-stoichiometric samples grown at 300°C.

obtain high-quality Fe_3Si/Ge structures with sharp interfaces without atomic mixing.

The cross-sectional TEM image of Fe_3Si/Ge grown at 130°C with the stoichiometric condition is shown in Fig. 9.12(a). This clearly demonstrates that the hetero-interface between Fe_3Si and Ge is atomically flat [44]. In addition, the electron diffraction pattern clearly indicates the characteristic spots due to fundamental reflection of Fe_3Si, superlattice reflection of B2-type Fe_3Si, and superlattice reflection of DO_3-type Fe_3Si. These results confirm that the Fe_3Si layers are mainly composed of the DO_3 ordering phase. The thermal stability of Fe_3Si formed at 60°C under stoichiometric and off-stoichiometric conditions was examined by post-annealing experiments (400°C for 30 min). Interdiffusion of Fe and Ge near the Fe_3Si/Ge interfaces is summarized in Figs 9.12(b) and (c), together with the χ_{min} values. Intermixing and deterioration of χ_{min} values are not found in the stoichiometric samples, which guarantees the thermal stability of Fe_3Si at interfaces up to 400°C [45]. However, in the off-stoichiometric samples, atomic mixing at the interface occurs significantly at 400°C, and the χ_{min} value increases to 48%, which suggests that excess Fe atoms over stoichiometric conditions cause the thermal stability of Fe_3Si/Ge structures to deteriorate. Consequently, accurate control of the Fe/Si ratio is very important in achieving high-quality Fe_3Si/Ge interfaces.

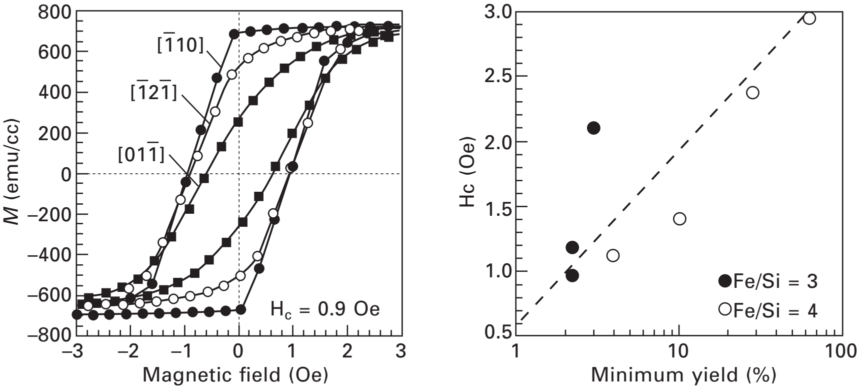
9.3.2 Magnetic and electrical properties of Fe_3Si/Ge, Si heterostructures

Magnetic hysteresis loops evaluated for Fe_3Si/Ge samples are shown in Fig. 9.13(a). The shapes of the loops depend on the field direction, indicating



(a) (b) (c)

9.12 (a) Cross-sectional TEM image of $Fe_3Si/Ge(111)$ grown at 130°C under stoichiometric condition; (b), (c) concentration profiles of Fe and Ge of (b) stoichiometric and (c) off-stoichiometric samples after annealing at 400°C.



(a) (b)

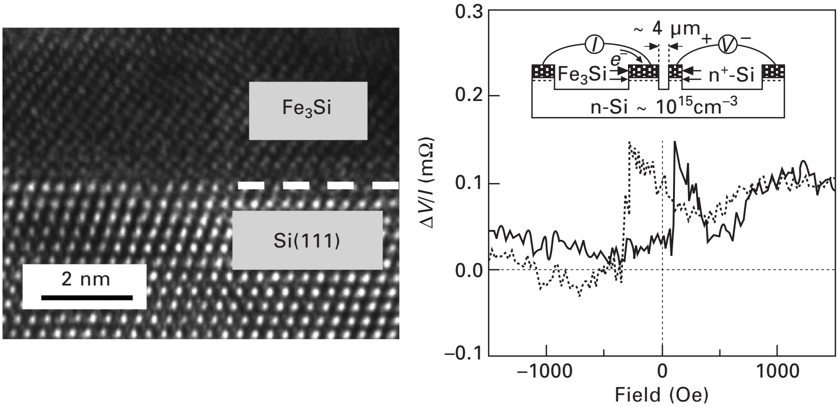
9.13 (a) Magnetic hysteresis loop of Fe_3Si layer (Fe:Si = 3:1, 130°C growth) for various field directions; (b) crystal quality (χ_{min}) dependent coercivities.

the presence of the magnetic anisotropy in the film plane. The coercivities were strongly dependent on the growth parameters, i.e., growth temperature and Fe/Si ratio. They are summarized in Fig. 9.13(b) as a function of χ_{min} evaluated by RBS measurements. The smallest value of the coercivity (0.9 Oe) is obtained from the sample with the highest crystal quality ($\chi_{min} = 2\%$). In addition, the anisotropy field was found to be very small (6 Oe) [48]. Consequently, formation of source/drain electrodes with uniform magnetic properties becomes possible by using the shape anisotropy.

Diode characteristics of Fe_3Si/Ge Schottky contacts evaluated by using current (*I*)–voltage (*V*) measurements indicated good Schottky characteristics with the barrier height of 0.56 eV, and a high ratio of the on-current to the off-current (~10^4). Post-annealing experiments guaranteed the thermal stability of such good electrical properties up to 400°C.

Heteroepitaxy of DO_3–Fe_3Si became possible even on Si(111) substrate by optimizing the growth parameters [46]. Cross-sectional TEM observation as shown in Fig. 9.14(a) clearly demonstrates the atomically flat interface between Fe_3Si and Si. Electrical properties evaluated by using I-V and C-V measurements indicate good Schottky characteristics with a barrier height of 0.62 eV and a high ratio of the on-current to the off-current (~10^4).

Electrical spin injection and detection were demonstrated by using the high-quality Fe_3Si/Si(111) interface [47]. Non-local voltage measurements [41, 52] were used to prove spin transport in a Si channel, where a δ-doped n^+-Si layer (~10^{19} cm^{-3}) was inserted between Fe_3Si and Si (~10^{15} cm^{-3}). A representative non-local spin signal ($\Delta V/I$) as a function of the external magnetic field is shown in Fig. 9.14(b). An evident spin-valve-like hysteretic feature at 2 K is shown. This proves that the spin-polarized electrons injected



(a) (b)

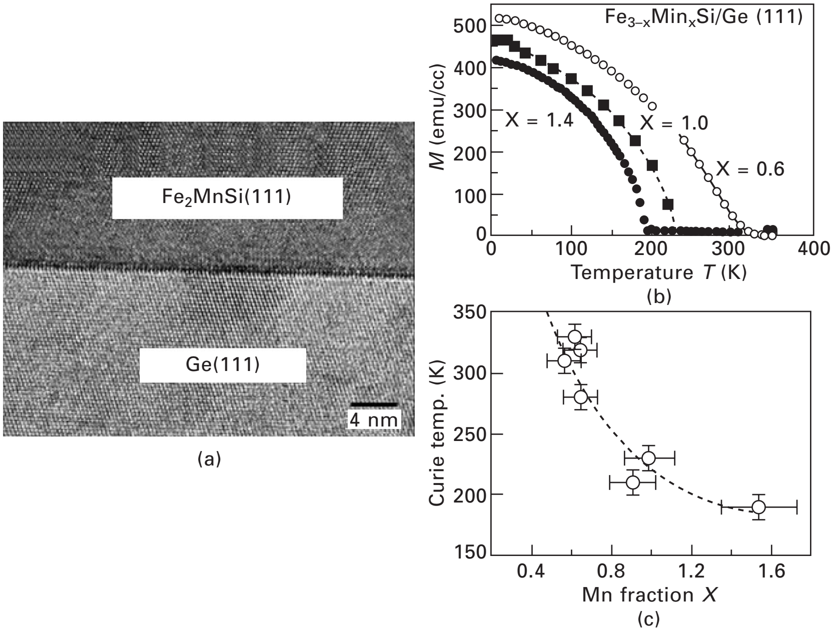
9.14 (a) Cross-sectional TEM image of Fe_3Si/Si Schottky contact; (b) a non-local spin signal through Fe_3Si/Si Schottky tunnel contact.

from an Fe_3Si contact are transported in the Si channel, and then the spins are detected by another Fe_3Si contact. Recently, spin injection and detection at a temperature (150 K) higher than that of liquid N_2 (77 K) was reported [53]. These results are now leading to expectations that spin transistors with SiGe channel will become a reality.

9.3.3 Atomically controlled epitaxial growth of full-Heusler alloys

Epitaxial growth (200°C) of a full-Heusler alloy, Fe_2MnSi, which has been expected to show half-metallic properties, on Ge(111) substrates has been reported [49, 54]. The cross-sectional TEM image is shown in Fig. 9.15(a), which clearly shows the formation of an atomically flat hetero-interface. In addition, the electron diffraction pattern of the grown Fe_2MnSi/Ge(111) indicated the formation of an ordered $L2_1$ phase, which is crucial in order to get half-metallicity. Magnetic field- and temperature-dependent magnetization curves showed a saturation magnetization (M_S) of ~2.17 μ_B/f.u. and a Curie temperature (T_C) of ~230 K. These results are similar to the magnetic properties reported in the bulk $Fe_{1.95}Mn_{1.05}Si$ sample [55].

The control of Mn concentration was examined to modify the magnetic properties and explore high-quality epilayers with room-temperature ferromagnetism [50]. Significant variations in the M_S and T_C were observed with keeping highly axial orientation, which was revealed by RBS measurements. The temperature dependence of the normalized magnetization for three samples ($x = 0.6$, 1.0, and 1.4) is shown in Fig. 9.15(b). For an off-stoichiometric composition of $x = 0.6$ or 1.4, T_C becomes higher or lower than 230 K,



9.15 (a) Cross-sectional TEM image of the grown Fe_2MnSi/Ge(111); (b) temperature dependence of normalized magnetization for three samples ($Fe_{3-x}Mn_xSi$: $x = 0.6, 1.0, 1.4$); (c) Mn fraction dependence of Curie temperature.

respectively. It should be noted that for $x = 0.6$, finite magnetic moments are observed at 300 K. The Curie temperature, summarized in Fig. 9.15(c), clearly indicates the ferromagnetic feature at room temperature for Fe-rich samples. The theoretical calculation predicts the half-metallicity of these samples [50].

Recently, another full-Heusler alloy, Co_2FeSi, with a high Curie temperature (1100 K) was successfully grown epitaxially on Ge(111) and Si(111) at low temperature (100°C) [51]. Such a low-temperature MBE technique will pave the way for future SiGe-based spintronics with epitaxial half-metallic materials.

9.4 Conclusion

Recent developments in miscellaneous crystal growth for SGOI hybrid structures and miscellaneous materials for SiGe spintronics are reviewed. Low-temperature crystallization of SiGe on insulating substrates becomes possible by SPC combined with ion beam irradiation, micro-indentation, and metal-imprint techniques. MILC using Ni as catalyst is a unique method for achieving SiGe crystals with large grains. In addition, application of electric fields during MILC enables directional nanowire growth aligned to the electric fields. Rapid melting growth initiated by SiGe mixing is a marvellous method which has achieved defect-free single-crystalline giant-Ge stripes on quartz substrates. High carrier mobility was demonstrated. SGOI hybrid structures obtained by these advanced methods open up the possibility of high-performance system-in-displays and 3-D LSIs.

Heteroepitaxy of ferromagnetic silicide (Fe_3Si) on SiGe becomes possible by using low-temperature MBE. By tuning the beam ratio (Fe/Si) and optimizing the growth temperature, Fe_3Si/SiGe structures with atomically flat interfaces were obtained. Reasonable magnetic properties comparable with bulk Fe_3Si and good Schottky characteristics were demonstrated. Such progress has enabled spin injection from the Fe_3Si electrode into the Si channel. In addition, heteroepitaxy of half-metallic alloys ($Fe_{2.4}Mn_{0.6}Si$, Co_2FeSi) on SiGe substrates was demonstrated. An evident ferromagnetic feature at room temperature was obtained. These results will be a powerful tool for realizing group-IV-semiconductor spin-transistors, consisting of Si and Ge channels with high mobility and a ferromagnetic source/drain for spin injection.

9.5 References

[1] M. Miyao and K. Nakagawa, *Jpn. J. Appl. Phys.* **33**, 3791 (1994).

[2] M. Miyao, E. Murakami, H. Etoh, K. Nakagawa, and A. Nishida, *J. Cryst. Growth* **111**, 912 (1991).

[3] M. Miyao, K. Nakagawa, Y. Kimura, and M. Hirao, *J. Vac. Technol. B*, **16**, 1529 (1998).

[4] T. Irisawa, S. Koh, K. Nakagawa, and Y. Shiraki, *J. Cryst. Growth* **215**, 670 (2003).

[5] J.S. Xia, Y. Ikegami, Y. Shiraki, N. Usami, and Y. Nakata, *Appl. Phys. Lett.* **89**, 201102 (2006).

[6] T. Mizuno, S. Takagi, N. Sugiyama, H. Satake, A. Kurobe, and A. Toriumi, *IEEE Electron Device Lett.* **21**, 230 (2000).

[7] M. Miyao, M. Tanaka, I. Tsunoda, T. Sadoh, T. Enokida, H. Hagino, M. Ninomiya, and M. Nakamae, *Appl. Phys. Lett.* **88**, 142105 (2006).

[8] N. Hirashita, Y. Moriyama, S. Nakaharai, T. Irisawa, N. Sugiyama, and S.Takagi, *Appl. Phys. Express* **1**, 101401 (2008).

[9] J.S. Im, H.J. Kim, and M.O. Thompson, *Appl. Phys. Lett.* **63**, 1969 (1993).

[10] C.H. Oh, M. Ozawa, and M. Matsumura, *Jpn. J. Appl. Phys.* **37**, L492 (1998).

[11] J. Olivares, A. Rodriguez, J. Sangrador, T. Rodriguez, C. Ballesteros, and A. Kling, *Thin Solid Films* **337**, 51 (1999).

[12] M. Miyasaka, K. Makihira, T. Asano, E. Polychroniadis, and J. Stoemenos, *Appl. Phys. Lett.* **80**, 944 (2002).

[13] M. Miyao, K. Toko, T. Tanaka, and T. Sadoh, *Appl. Phys. Lett.* **95**, 022115 (2009).

[14] S. Datta and B. Das, *Appl. Phys. Lett.* **56**, 665 (1990).

[15] S. Sugahara and M. Tanaka, *Appl. Phys. Lett.* **84**, 2307 (2004).

[16] K. Hamaya, S. Masubuchi, M. Kawamura, T. Machida, M. Jung, K. Shibata, K. Hirakawa, T. Taniyama, S. Ishida, and Y. Arakawa, *Appl. Phys. Lett.* **90**, 053108 (2007).

[17] Y. Ando, K. Hamaya, K. Kasahara, Y. Kishi, K. Ueda, K. Sawano, T. Sadoh, and M. Miyao, *Appl. Phys. Lett.* **94**, 182105 (2009).

[18] T. Sadoh, H. Kamizuru, A. Kenjo, and M. Miyao, *Appl. Phys. Lett.* **89**, 192114 (2006).

[19] S. Yamaguchi, N. Sugii, K. Park, K. Nakagawa, and M. Miyao, *J. Appl. Phys.* **89**, 2091 (2001).

[20] I. Tsunoda, T. Nagata, A. Kenjo, T. Sadoh, and M. Miyao, *Mat. Sci. Eng. B* **89**, 336 (2002).

[21] I. Tsunoda, K. Nagatomo, A. Kenjo, T. Sadoh, S. Yamaguchi, and M. Miyao, *Jpn. J. Appl. Phys.* **43**, 1901 (2004).

[22] M. Miyao, I. Tsunoda, T. Sadoh, and A. Miyauchi, *J. Appl. Phys.* **97**, 4909 (2005).

[23] H. Kanno, I. Tsunoda, A. Kenjo, T. Sadoh, and M. Miyao, *Appl. Phys. Lett.* **82**, 2148 (2003).

[24] H. Kanno, K. Toko, T. Sadoh, and M. Miyao, *Appl. Phys. Lett.* **89**, 182120 (2006).

[25] H. Kanno, A. Kenjo, T. Sadoh, and M. Miyao, *Jpn. J. Appl. Phys.* **45**, 4370 (2005).

[26] D.K. Choi, H.-C. Kim, and Y.-B. Kim, *Appl. Phys. Lett.* **87**, 063108 (2005).

[27] N. Higashi, G. Nakagawa, T. Asano, M. Miyasaka, and J. Stoemenos, *Jpn. J. Appl. Phys.* **45**, 5B, 4347 (2006).

[28] H. Watakabe, T. Sameshima, H. Kanno, T. Sadoh, and M. Miyao, *J. Appl. Phys.* **95**, 6457 (2004).

[29] M. Miyao, T. Tanaka, K. Toko, and M. Tanaka, *Appl. Phys. Express* **2**, 045503 (2009).

[30] K. Toko, T. Sakane, T. Tanaka, T. Sadoh, and M. Miyao, *Appl. Phys. Lett.* **95**, 112107 (2009).

[31] M. Kurosawa, N. Kawabata, T. Sadoh, and M. Miyao, *Appl. Phys. Lett.* **95**, 132103 (2009).

[32] Y. Liu, M.D. Deal, and D. Plummer, *Appl. Phys. Lett.* **84**, 2563 (2004).

[33] D.J. Tweet, J.J. Lee, J.S. Maa, and S.T. Hsu, *Appl. Phys. Lett.* **87**, 141908 (2005).

[34] J. Feng, G. Thareja, M. Kobayashi, S. Chen, A. Poon, Y. Bai, P.B. Griffin, S.S. Wong, Y. Nishi, and J.D. Plummer, *IEEE Electron Device Lett.* **29**, 7, 805 (2008).

[35] K. Toko, T. Sadoh, and M. Miyao, *Appl. Phys. Lett.* **94**, 192106 (2009).

[36] K. Toko, H. Kanno, A. Kenjo, T. Sadoh, T. Asano, and M. Miyao, *Appl. Phys. Lett.* **91**, 042111 (2007).

[37] J. Jang, J.Y. Oh, S.K. Kim, Y.J. Choi, S.Y. Yoon, and C.O. Kim, *Nature* **395**, 481 (1998).

[38] M. Tamura, H. Tamura, M. Miyao, and T. Tokuyama, *Proc. 12th Conf. Solid State Devices,* Tokyo, 1980, *Jpn. J. Appl. Phys.* **20**, Suppl. 20-1, 43 (1981).

[39] Y. Maeda, K.P. Homewood, T. Suemasu, T. Sadoh, H. Udono, and K. Yamaguchi, *Thin Solid Films* **461**, 1 (2004); Y. Maeda, K.P. Homewood, T. Sadoh, Y. Terai, K. Yamaguchi, and K. Akiyama, *Thin Solid Films* **515**, 8101 (2007).

[40] S.A. Wolf, D.D. Awschalom, R.A. Buhrman, J.M. Daughton, S. von Molnar, M.L. Roukes, A.Y. Chtchelkanova, and D.M. Treger, *Science* **294**, 1488 (2001).

[41] B. Huang, D.J. Monsma, and I. Appelbaum, *Phys. Rev. Lett.* **99**, 177209 (2007).

[42] O.M.J. van't Erve, A.T. Hanbicki, M. Holub, C.H. Li, C. Awo-Affouda, P.E. Thompson, and B.T. Jonker, *Appl. Phys. Lett.* **91**, 212109 (2007).

[43] J. Herfort, H.-P. Schönherr, and K.H. Ploog, *Appl. Phys. Lett.* **83**, 3912 (2003).

[44] T. Sadoh, M. Kumano, R. Kizuka, K. Ueda, A. Kenjo, and M. Miyao, *Appl. Phys. Lett.* **89**, 182511 (2006).

[45] Y. Maeda, T. Jonishi, K. Narumi, Y. Ando, K. Ueda, M. Kumano, T. Sadoh, and M. Miyao, *Appl. Phys. Lett.* **91**, 171910 (2007).

[46] K. Hamaya, K. Ueda, Y. Kishi, Y. Ando, T. Sadoh, and M. Miyao, *Appl. Phys. Lett.* **93**, 132117 (2008).

[47] Y. Ando, K. Hamaya, Y. Kishi, K. Kasahara, K. Ueda, K. Sawano, T. Sadoh, and M. Miyao, *Appl. Phys. Lett.* **94**, 182105 (2009).

[48] Y. Ando, K. Hamaya, K. Kasahara, K. Ueda, Y. Nozaki, T. Sadoh, Y. Maeda, K.Matsuyama, and M. Miyao, *J. Appl. Phys.* **105**, 07B102 (2009).

[49] K. Ueda, K. Hamaya, K. Yamamoto, Y. Ando, T. Sadoh, Y. Maeda, and M. Miyao, *Appl. Phys. Lett.* **93**, 112108 (2008).

[50] K. Hamaya, H. Itoh, O. Nakatsuka, K. Ueda, K. Yamamoto, M. Itakura, T. Taniyama, T. Ono, and M. Miyao, *Phys. Rev. Lett.* **102**, 137204 (2009).

[51] S. Yamada, K. Hamaya, K. Yamamoto, T. Murakami, K. Mibu, and M. Miyao, *Appl. Phys. Lett.* **96**, 082511 (2010).

[52] F.J. Jedema, H.B. Heersche, A.T. Filip, J.J.A. Baselmans, and B.J. van Wees, *Nature (London)* **416**, 713 (2002).

[53] Y. Ando, K. Kasahara, Y. Enomoto, K. Yamane, K. Hamaya, K. Sawano, T. Kimura, and M. Miyao, *11th Joint MMM-INTERMAG Conf.*, 18–22 January 2010, Washington, DC, BE-03, 340.

[54] S. Fujii, S. Ishida, and S. Asano, *J. Phys. Soc. Jpn* **64**, 185 (1995).

[55] S. Yoon and J.G. Booth, *J. Phys. F: Metal Phys.* **7**, 1079 (1977).

10

Modeling the evolution of germanium islands on silicon(001) thin films

L. MIGLIO and F. MONTALENTI, University of Milano-Bicocca, Italy

Abstract: Ge/Si is popularly considered as the prototypical example of a system following the Stransky–Krastanow (SK) growth modality. Despite being truly simpler than other SK systems, Ge/Si displays remarkably fascinating physics, starting from the very onset of island formation, discussed in detail in this chapter with the help of a combined theoretical and experimental analysis. Complex phenomena following nucleation of 3D structures, such as Si/Ge intermixing and island ordering, are also described and interpreted with the use of simple theory. Some recent findings, likely to trigger further research in the coming years, are finally summarized in the concluding future trends section.

Key words: semiconductor thin-films, Stransky–Krastanow growth, Ge/Si, strain release, nanoislands.

10.1 A few considerations on epitaxial growth modeling

Ge on Si(001) is popularly considered as a prototypical system to understand the physics of Stranski–Krastanow (SK) growth, which also characterizes other more complex systems of greater applicative interest such as InGaAs/GaAs(001). As we shall see, prototypical does not mean simple. Ge growth on Si(001) is indeed a multiscale problem, where atomic-scale information (in terms, e.g., of details of the surface reconstructions, or kinetic barriers) is a key ingredient in order to predict properties on much longer scales. The typical island dimensions (up to hundreds of nanometers in diameter), and the long-range nature of the elastic field, make Ge/Si a system for which multiple theoretical approaches must be mixed together. If density functional theory (DFT) seems to be the only approach able to quantify the relative stability of the different facets which are exposed during Ge/Si(001) SK growth, or adatom/ad-dimer activation energies for diffusion, computational limitations prevent using DFT to evaluate the strain field of large nanostructures (some 100 nm in size) on top of an interacting thick substrate. Classical molecular dynamics (MD) approaches and (more suitably) continuum elasticity theory implemented by finite element methods (FEM) are valid alternatives, although some care is needed when using, in one model, parameters determined by different approaches. It is interesting to note, as we directly verified for several island geometries, that FEM is a reliable tool down to a few nanometers if one is willing to accept small deviations from linearity.

The formation and evolution of epitaxial islands can be investigated either by kinetic models or by thermodynamic ones. Both approaches suffer from some limitations. While in principle any crystal growth study should be performed including kinetics, this leads unavoidably to strong simplifications, e.g. in the description of the system geometry, of the diffusion processes, and/or of the strain fields. Realistic modeling of heteroepitaxial growth kinetics at the atomic scale (e.g. by kinetic Monte Carlo simulations) is truly a daunting task, if one considers that on complex reconstructed surfaces subject to strain, a single adatom does not necessarily reflect the system thermodynamics [1], and the energetics of a newly formed layer change qualitatively as small atomic aggregates form [2].

On general grounds, the modifications in shape and arrangement of the islands with Ge deposition at the common experimental temperatures (600–700°C) are essentially driven by a flow of surface atoms, moving and resting on a multifaceted landscape, according to the difference in chemical potential $\mu(x, y)$. The latter is modulated both by the surface energies of incomplete, maybe terraced facets, and by the in-plane strain components produced by the composition and the morphology below the flowing layer. In order to gather a qualitative understanding of the driving forces influencing atomic motion and aggregation at least at an average level, a physically sound starting point of view is to consider an ideal 'disc' of Ge atoms probing the variations of the local in-plane lattice parameter and average surface energies, which – in turn – corresponds to variations in $\mu(x, y)$. This corresponds to a 'local' thermodynamic analysis of kinetic issues, which is not a novelty in materials science, as the standard nucleation theory itself is based on a similar approach.

On the other hand, the fully thermodynamic approach, i.e. the unrestricted search for the system configuration of the lowest free energy, is not applicable here. In typical experiments, only a few layers of Ge are deposited on thick Si buffer layers. The thermodynamic limit would correspond to having Ge diluted into the substrate, destroying the main driving force for SK growth, i.e. strain relaxation. This limit is never reached since bulk diffusion is not fully active in the typical experimental conditions [3]. Any reliable thermodynamic model should always consider this limitation: the only relevant portion of the phase space which is visited by the system is the one reachable by surface diffusion only. The extent of such a condition is often imposed using the help of the experimental evidence: how deep is the reshuffling of the surface (flowing) layer, especially when Si atoms are popping up from the substrate? This is still an open and intriguing question.

However, in this chapter we shall see examples where 'restricted' thermodynamics, variably guided by experiments, provides a nice understanding of SK growth, helping in unraveling complex experimental evidence.

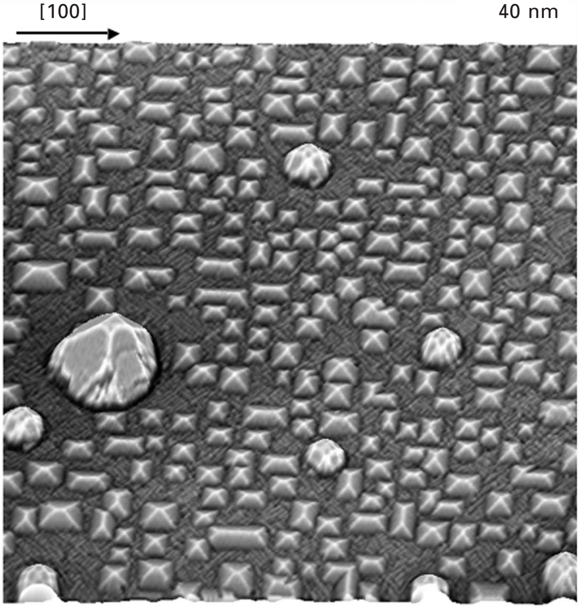
10.2 Introduction to Stranski–Krastanow (SK) heteroepitaxy

Let us consider a substrate of material A on top of which atoms of material B are deposited, A and B being lattice mismatched, such as Ge and Si, or InGaAs and GaAs. We say that the heteroepitaxial growth of B on A follows the SK modality if at low coverages B wets A, forming a pseudomorphic wetting layer (WL), while at later deposition stages three-dimensional B islands are formed. Formation of 3D structures in SK growth is driven by thermodynamics, being a direct manifestation of the attempt of the B material to recover its bulk lattice parameter, releasing at least part of the elastic energy caused by the lattice mismatch. As such, SK islands should not be confused with kinetically induced ones, e.g. produced by layer-by-layer breakdown owed to large Ehrlich–Schwoebel barriers [4]. Finally, the SK growth differs from the so-called 'Volmer–Weber' (VW) modality, where islands are formed straightforwardly, without interposing a WL, in that the former includes a sudden transition between the two-dimensional growth and the three-dimensional one, which is still an open fundamental aspect of heteroepitaxy. The standard 'capillarity model' [5] nicely explains the different growth modalities (layer-by-layer, named Frank–van der Merwe (FM), SK and VW) on the basis of total surface energy minimization, including that of the substrate, that of the film (strain energy included) and the interface energy of the substrate with the film. However, in realistic SK systems, the morphological evolution of the islands with deposition cannot be captured by such a simple model and a more detailed approach is needed. Indeed, it is the competition between the actual surface energy of the island facets and the elastic energy relaxation with increasing aspect ratio (height to base ratio) which determines the evolution of the island shape with volume.

Let us consider here the Ge/Si(001) system, which is a prototypical one for SK growth, due to the absence of bond polarity effects and the full miscibility of the Ge–Si systems. Islanding in a Ge/Si(001) system was first evidenced in 1990 [6, 7]. A great amount of experimental work and modeling interpretations followed in the next 20 years (reviews can be found in [8]–[12]), reporting on various island shapes. The shallowest islands are called 'prepyramids' or 'mounds' [13, 14] and display an aspect ratio (ar) lower than 0.1, while slightly steeper are {105} square-based pyramids (ar = 0.1) [6, 15] or rectangular huts (obtained instead of pyramids at sufficiently low temperatures). Multifaceted 'domes' (ar = 0.22–0.26) reflect the appearance of {113} facets [15, 16] and others [17], and were thought to be the ultimate evolution stage before nucleation of misfit dislocations, until steeper (ar = 0.26–0.32) 'barn' morphologies [18] were seen. In Fig. 10.1, a nice STM picture taken from [19] shows many such shapes being simultaneously present on the same sample.

For many years attention was focused mainly on {105} pyramids and domes. It was pointed out how in the initial stages of island formation these two morphologies coexist, pyramids characterizing lower-volume islands [15]. As clarified by Ross *et al.* [20], such pyramid–dome bimodal distribution characterizes a well-defined time interval accompanying deposition. At long exposure times (high coverages), only steeper islands survive, giving rise to an almost unimodal distribution of domes. By proceeding with deposition, however, large dislocated islands start coexisting with domes or barns, until such 'superdomes' (largest island in Fig. 10.1) dominate the population [21–23].

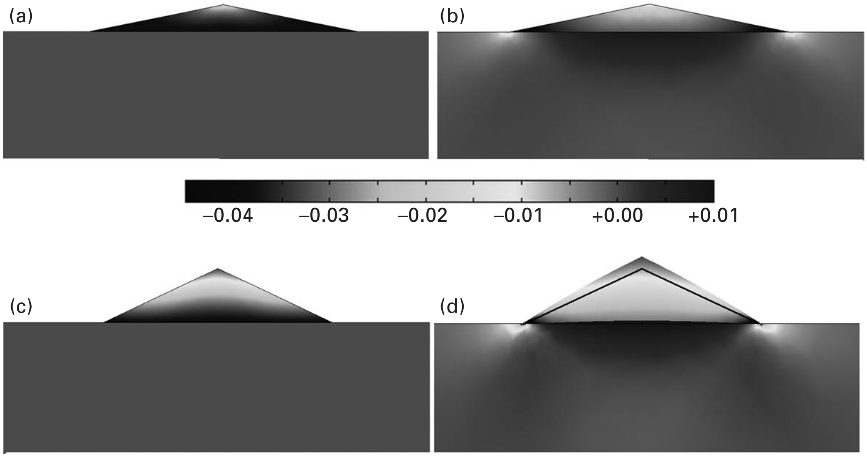
Although we investigated in recent years the plastic relaxation mechanism [24, 25], eventually providing an intriguing path to material accumulation



10.1 STM image showing formation of Ge islands of various shapes on Si(001), coexisting on the same sample. 7 ML of Ge were deposited on Si(001) by magnetron sputtering. The rather low growth temperature (500°C) allows for the coexistence not only of square-basis (105) pyramids and rounded domes, but also of rectangular (105) huts. The figure also shows the presence of a single dislocated island (larger island, left portion of the figure). From [19], reprinted with the author's permission.

in the islands by the ‘cyclic growth’ [21], in this chapter we shall limit our interest to the elastic relaxation mechanisms, particularly important for the early stages of growth, i.e. the SK onset.

In-plane compressive strain in Ge (or SiGe) islands originates from lattice misfit, which is as large as 4.2% in the case of pure Ge on Si. Correspondingly, the strain component perpendicular to the substrate is positive (i.e. extensive), according to the Poisson ratio, aiming at preserving the cell volume. This is the actual situation for continuous Ge (GeSi) films on Si(001), i.e. for the wetting layer. For three-dimensional islands, instead, the physics involved in the relaxation mechanism is more complex, and we shall describe it with the help of Fig. 10.2, where FEM results for the ε_{xx} component of the strain tensor (x being the horizontal direction) are reported for a shallow (ar = 0.1, corresponding to {105} facets) and a steeper (ar = 0.235, corresponding to {113} facets) Ge pyramid on a Si substrate. Elastic relaxation occurs essentially by three mechanisms. The first one, acting at any aspect ratio, even for very flat islands, is the strain redistribution from the island to the substrate. The second is the free expansion at the island facets [26], inducing the tendency towards recovering the correct lattice parameter at the island top. To separate the above effects it is useful to compare two different calculations, one performed by imposing a frozen substrate (Fig.



10.2 (a), (b) Strain maps (ε_{xx} component of the strain tensor, x being the horizontal direction) for a shallow Ge pyramid (ar = 0.1) on Si as obtained by 3D FEM calculations (only a transversal section is shown). In (a) the Si substrate is kept frozen at its lattice parameter, while substrate relaxation is allowed in (b). In (c) (with frozen substrate) and (d) calculations are repeated for a steeper island (ar = 0.235). In (d) the map is shown using a deformation plot, highlighting the island shape change (artificially raised by a factor of 3).

10.2(a) for the shallow pyramid, Fig. 10.2(c) for the steeper one), the other allowing for island and substrate equilibration (Figs 10.2(b) and (d)). From the FEM results it is clear that island/substrate strain redistribution is the main mechanism for low aspect-ratio islands. By keeping the Si underneath frozen, indeed, only a tiny relaxation effect is seen in Fig. 10.2(a), while the whole top half of the pyramid appears sizably relaxed when substrate relaxation is considered. Free expansion at the island facets, instead, becomes sizeable only for aspect ratios larger than some 0.2 (Fig. 10.2(c)), while it is sometimes erroneously believed to be the leading mechanism at any aspect ratio (see the discussion in [27]).

The cumulative effect of strain relaxation is a compressive load close to the island base perimeter (the black region in the island lower periphery in Figs 10.2(b) and (d)), and in the substrate below it (the light lobes in the substrate in Figs 10.2(b) and (d)), an in-plane relaxation toward the island apex (growing with the aspect ratio – compare Figs 10.2(b) and (d)), and an expansion (tensile strain – the dark region below the island in Figs 10.2(b) and (d)) in the substrate vertically below it. Overall, island relaxation is accompanied also by a shear deformation (also involving the substrate), resulting in the deformed shape displayed in Fig. 10.2(d). A more detailed analysis of the strain tensor for islands of different aspect ratio can be found in [28].

Despite being somewhat artificial, the above-proposed description of strain relaxation in terms of two different mechanisms is conceptually convenient. For example, the idea that for small aspect ratio islands the key contribution in lowering the effective lattice misfit is provided by a load redistribution between island and substrate is the basis of the popular 'flat-island approximation [29]', allowing for quick estimates of strain release in shallow islands. Also, the same mechanism allows one to understand the peculiar relaxation taking place on pit-patterned substrates, where the low energetic budget [30] of the Ge initially filling the pit bottom (exposing a flat (001) facet) can be explained purely by load redistribution, since Ge does not form inclined free facets [31].

Finally, a third relaxation mechanism, strongly dependent on the growth conditions, takes place. The high compression at the island base perimeter (see Fig. 10.2 and Refs [32, 33]) facilitates the creation of trenches expelling both Ge atoms from the wetting layer and Si atoms from the substrate, the latter intermixing with the Ge (therefore increasing the entropy of mixing [34, 35]) and progressively accumulating at the island facets, particularly in the lower part [36]. This intermixing generates an inhomogeneous composition (see Section 10.4), the average Ge content strongly decreasing with temperature [37–41].

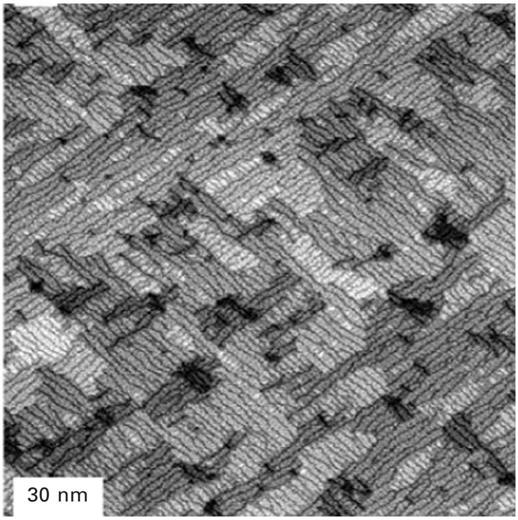
Island surface energies increase with exposed surface area, i.e. with aspect ratio, and usually with increasing Miller indexes of the facets, but there are sizeable exceptions to this rule of thumb, as produced by smart surface reconstructions. One example is the rebonded step reconstruction of the (105) surface, which stabilizes the pyramid shape [42] and very frequently appears in different Ge island shapes and film morphologies [43, 44] due to the increasing stability with compressive strain [45–47].

Therefore, the thermodynamic origin of the morphological evolution with Ge deposition is apparently well established. A flat wetting layer as thick as a few monolayers (ML) is first produced, in order to decrease the surface energy of Si(001) to its value for Ge(001). Then, by further deposition, the increasing elastic energy budget originated by the lattice misfit generates the appearance of three-dimensional islands in the form of square-based pyramids or rectangular-based huts with (105) facets, frequently preceded by instabilities in the flat film, such as mounds or pre-pyramids. Actually, such shallow pyramids allow for a limited amount of strain to be transferred to the substrate [27] and the rebonded-step reconstruction of the 105 facets is needed in stabilizing this shape [42, 45]. As deposition of Ge goes on, the size of the island increases and the limited efficiency of the shallow pyramids in relieving the compressive stress drives a progressive shape evolution to islands of higher aspect ratios, such as domes [15, 16, 48, 49] and barns [18]. Here, the expansion of the lattice component parallel to the surface occurring close to the free facets overcompensates for the larger surface energy of the steeper facets and a sizeable SiGe intermixing is produced by the spillout of Si atoms from the trenches originating at the base of the domes [37–41], in turn decreasing the elastic energy budget.

We have discussed the appearance of islands of different shapes, considering the further complexity brought in by SiGe intermixing. Let us now focus on the very initial stages of growth. Understanding the onset of SK growth is particularly hard owing to the abrupt 2D/3D transition [20], making it difficult to analyze intermediate states between a flat WL and the first 'real' faceted islands, {105} pyramids. Indeed, the role played by the WL during SK onset was fully revealed in its complexity only recently [50].

When Ge is deposited on a dimerized Si(001) (1 × 2) surface, strain destabilizes the original reconstruction so that after depositing ~1 ML a $(2 \times N)$ symmetry is observed. This is caused by the interruption of dimer rows by vacancy lines. Between ~1 ML and ~2.5 ML the periodicity changes, vacancy lines getting progressively closer, and N reaching a minimum value of 6–9, depending on the growth conditions. Proceeding with deposition, dimer vacancy rows also appear, leading to WL breaking into small patches as shown in the STM image of Fig. 10.3. This patched structure is often called $(M \times N)$ reconstruction, although the value of the M-periodicity seems to be defined only as an average, and the distribution of terrace sizes is quite broad [9, 51–53].

Missing-dimer configurations act as small trenches, relieving part of the



10.3 ($M \times N$) Ge/Si(001) reconstruction for a WL as thick as 4 ML, grown by magnetron sputtering at 450°C. In the figure, dimer rows run from the bottom right corner to the top left one. From [53], reprinted with the authors' permission. Copyright (2003) by the American Physical Society.

compressive strain felt by Ge atoms. DFT calculations proved to be important in characterizing and understanding the WL surface energy [46, 53–55]. In particular, Lu *et al.* [46] found the (2×10) reconstruction to be the favored one below 1 ML, while the (2×8) is preferred for further deposition. The semi-quantitative agreement with the experiments demonstrates that the observed changes in reconstruction are driven by thermodynamics. Importantly, DFT calculations also showed the dependence of the WL surface energy γ_{WL} on the thickness of the WL. From [46, 55] it is clear that γ_{WL} decreases with the number of Ge layers placed on Si(001), and it takes some five layers before convergence is reached, γ_{WL} approaching the surface energy of Ge(001), tetragonally strained at the Si/Ge misfit (the small Ge/Si interfacial can often be neglected [55]). While a consistent drop in γ_{WL} is expected when a first layer of Ge is placed on Si, the persisting decrease at higher coverages is more intriguing, signaling how the outermost Ge layers still feel the presence of the Si substrate underneath, influencing (raising) the Ge dangling-bond energy [54]. The decrease in γ_{WL} with the thickness can be equivalently [50, 55] analyzed in terms of an increase of the average chemical potential μ of the outermost Ge layer. For one and two layers the DFT calculations of Beck *et al.* [55] show that the value of μ is actually lower than μ for Ge bulk. This means that only for coverages exceeding 2 ML can one expect Ge atoms to eventually arrange themselves in 3D islands to release the strain. Before, energy minimization drives the formation of a flat WL due to the electronic effect discussed above. Interestingly, also a classical model predicts γ_{WL} to decrease with thickness, due to the different influence of the stress induced by the surface reconstruction on the underlying layers [56]. Elastic effects, thus, also influence the γ_{WL} behavior. As in the case of islands, also the WL can host non-negligible amounts of Si. Surface atomic Ge/Si exchange events take place in the experimental temperature range, as revealed by DFT calculations and experiments [3, 57, 58]. The Ge concentration in the WL can be estimated by looking at confinement shifts in photoluminescence (PL) spectra. From recent results, it appears that intermixing in the WL is less pronounced than in islands. For instance, at $T = 700$°C a Ge content of $c \sim 0.85$ was revealed. Islands appearing at later deposition stages, instead, displayed only $c \sim 0.35$ [50, 59].

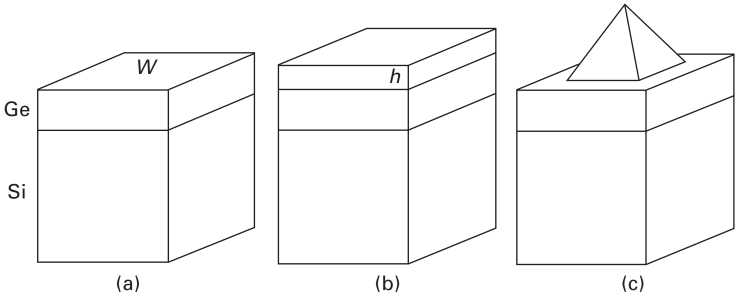
If Beck *et al.* [55] demonstrated that formation of stable Ge islands is not possible below a coverage of at least 2 ML, for further deposition elastic relaxation can compete with WL thickening. Let us analyze this issue, with the help of a simple model.

10.3 Onset of Stranski–Krastanow (SK) heteroepitaxy

10.3.1 A simple thermodynamic model for SK growth

A model of SK growth aimed at understanding when 3D island formation is favored with respect to WL thickening can be formulated in the following terms. We start by considering a Si substrate covered by N Ge layers (N-WL in the following), and we ask whether it is energetically more convenient to add N_{at} additional Ge atoms as a new layer of the WL, or to arrange them into a 3D island, placed over the N-WL. For simplicity in the following we shall consider only volumetric and surface contributions to the total energy of the system, therefore neglecting step and/or edge energies [8]. This allows us to avoid treating terms linked to the finite areal extension of the WL, simply considering the situation schematized in Fig. 10.4.

If N is sufficiently large, i.e. we are in the thick-film limit where γ_{WL} no longer depends on N, then the $(N + 1)$-WL (Fig. 10.4(b)) and the N-WL have the same surface energy, while the elastic energy differs by $V\rho_{\mathrm{t}}$, ρ_{t} being the volumetric density of elastic energy of Ge, tetragonally strained due to the Si/Ge misfit, and V the volume occupied by the N_{at} additional Ge atoms. If, instead, we compare the N-WL with the island configuration (Fig. 10.4(c)), the difference in elastic energy is $V\rho_{\mathrm{is}}$, where ρ_{is} is the volumetric density of elastic energy in the island. Here, we are assuming that N_{at} Ge atoms occupy the same volume V in both the WL and island configurations. This is not exact, but corrections are rather small [50]. ρ_{is} depends on the



10.4 (a) A Ge WL made of *N* Ge layers is deposited over a Si substrate. (b) A further single layer of Ge is added to the WL, now made of (*N* + 1) layers. (c) An island is formed over the original WL. The volume of the island is *V* = *Wh*, *W* being the areal extension of the *WL* and *h* the height of a single Ge layer.

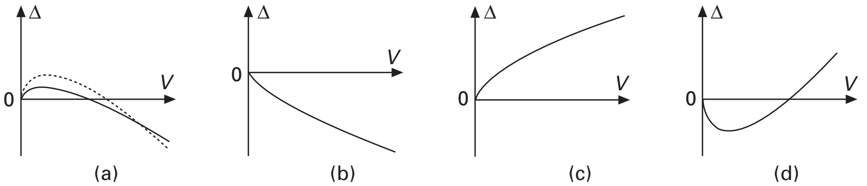
island shape, and includes the substrate deformation induced by the presence of a 3D structure (see Section 10.3.2). Furthermore, the island and WL configurations have different surface energy. If S is the extension of the island's free facets, and A the WL covered by the island base, the surface energy difference is $S\gamma_{\text{is}}$, $-A\gamma_{\text{WL}}$, where γ_{is} is the average surface energy (density) of the islands' facets, averaged over the facets and over strain, as discussed in Section 10.3.2. The energy difference between the island and the (N+1)-WL is therefore

$$\Delta = V(\rho_{\text{is}} - \rho_{\text{t}}) + S\gamma_{\text{is}} - A\gamma_{\text{WL}} \qquad 10.1$$

If, for a given V, $\Delta < 0$, formation of an island at that volume is favored with respect to WL thickening. For an island growing self-similarly, $S = sV^{2/3}$, and $B = bV^{2/3}$, s and b being geometrical parameters depending on the island shape. Therefore

$$\Delta(V) = DV + EV^{2/3} \qquad 10.2$$

where $D = (\rho_{\text{is}} - \rho_{\text{t}})$, and $E = (s\gamma_{\text{is}} - b\gamma_{\text{WL}})$. For large enough volumes, the sign of Δ is determined solely by D. Since any of the islands' shapes discussed in Section 10.2 allow for a better elastic relaxation with respect to the simple tetragonal one, $D < 0$. In this thick-film limit all islands are characterized by a finite critical volume, beyond which their formation is favored by thermodynamics. Furthermore, if $\gamma_{\text{is}} \approx \gamma_{\text{WL}}$ then $E > 0$, so that the $\Delta = \Delta(V)$ curve has the shape displayed in Fig. 10.5(a). At small volumes, the energy cost in terms of extra-exposed surface favors the WL configuration, while at larger V values strain relaxation dominates, and thermodynamics drives island formation and enlargement. Notice that islands tend to enlarge also for volumes lower than the critical one (here defined as V: $\Delta(V) = 0$), provided that the maximum of the curve is reached. From there, indeed, the difference



10.5 Energy difference between a 3D island and the WL as a function of the island volume. In (a) the volumetric term of Eq. 10.2 is negative ($D < 0$), while the surface one is positive ($E > 0$). Island formation is driven by thermodynamics only above a critical volume. Also, the dotted curve represents a steeper island with respect to the one represented by a continuous line, under the assumption of similar surface energy densities (see Section 10.3.2). (b) Island formation is favored by both the volumetric and the surface term, leading to barrierless nucleation (null critical volume). (c) Formation of stable islands in never possible, due to positive volumetric and surface terms. (d) The volumetric term is positive, while the surface one is negative. Thermodynamics drives island formation only at small volumes.

in chemical potential ($\approx \partial\Delta/\partial V$) between island and WL becomes negative. The curve displayed in Fig. 10.5(a) is typical of nucleation processes, and displays a barrier to be overcome before islands can form and grow.

Formally, it is simple to extend the model including the effect of an N-dependent wetting layer surface energy $\gamma_{\mathrm{WL}}(N)$. As shown in [50], indeed, it is sufficient to replace Eq. (10.1) with the slightly more complex expression for Δ:

$$\Delta = V\,[\rho_{\mathrm{is}} - \rho_{\mathrm{eff}}\,(N+1)] + S\gamma_{\mathrm{is}} - A\gamma_{\mathrm{WL}}(N) \qquad 10.3$$

where $\rho_{\mathrm{eff}}\,(N+1) = \rho_{\mathrm{t}} + \frac{1}{h}\,[\gamma_{\mathrm{WL}}\,(N+1) - \gamma_{\mathrm{WL}}\,(N)]$, h representing the height of a single Ge layer in the WL, and $\gamma_{\mathrm{WL}}(N)$ is the surface energy of N layers of Ge deposited on Si(001). Comparing Eq. 10.3 with Eq. 10.1, the main difference is clear. From $\gamma(N+1) < \gamma(N)$ [46, 55], it follows that $\rho_{\mathrm{eff}}\,(N+1) < \rho^{\mathrm{t}}$, meaning that the volumetric term is negative (driving island formation) only if strain relaxation is sufficiently strong. In other words, for a fixed N it is possible that some islands become stable beyond a finite critical volume, while others do not.

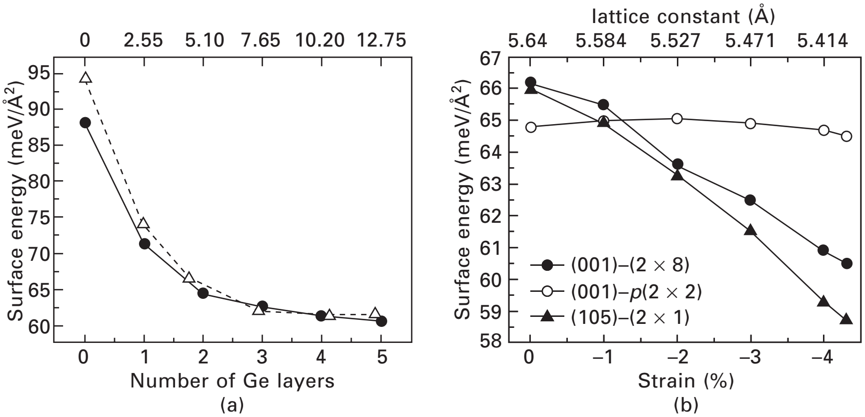
Let us now show how the various terms entering Eq. (10.3) can be quantified, leading to a possible comparison with experimental data.

10.3.2 Surface and volumetric terms in Ge/Si(001)

Equations similar to Eq. 10.3 have been discussed in general terms in the literature (see, e.g., [60]), showing a wealth of different behaviors depending on the various parameters. Using a different perspective, we instead wish to particularize Eq. 10.3 to Ge/Si(001) by limiting the number of adjustable parameters.

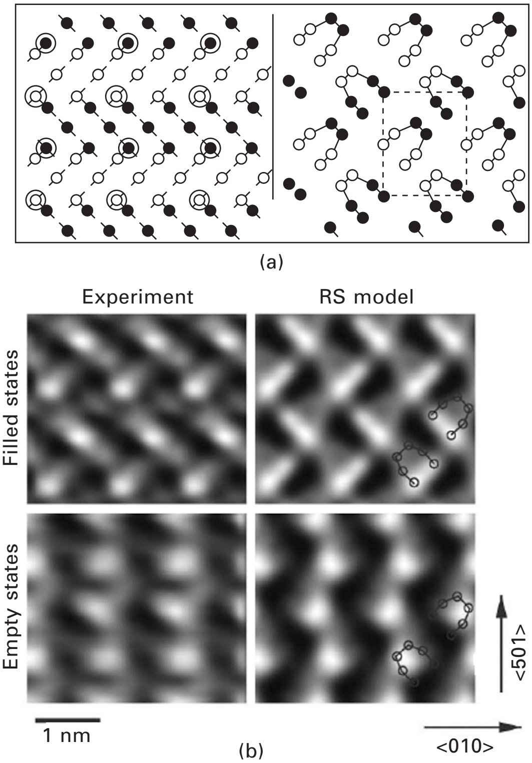
Let us first consider surface energy terms which require atomistic approaches to be evaluated, since they are directly related to bond breaking, bond distortion and rehybridization. While classical potentials have been shown to give important qualitative information [45, 56], only first-principles approaches can tackle the complex energetics involved whenever the simple sp^3 hybridization is lost. Concerning the surface energy $\gamma_{WL}(N)$ of the WL, this has been evaluated by DFT as a function of the WL layers N and also of strain [46, 47, 55, 61]. An example is given in Fig. 10.6, taken from [46].

Computing by DFT the surface energy of the exposed facets is more complex. Due to computational limitations, indeed (but we signal a recent methodological development that could allow for more direct calculations [62]), γ for an {l,m,n} facet is typically computed using a planar slab terminated with the {l,m,n} orientation of the facet. Beside problems linked to the small extension of the facet (negligible for large enough islands), the main difficulty stems in considering the strain-dependent behavior of γ, which was shown to be rather important for both (001) and (105) (see Fig. 10.6). For simplicity, indeed, this is taken to be biaxial, an assumption which is exact for (001) and probably an excellent approximation for (105) but could become increasingly incorrect for steeper facets. In any case, strain-dependent calculations have focused so far only on (001) and (105)



10.6 (a) Filled circles: surface energy of Ge/Si(001) as a function of the number of Ge layers; for each N, the lowest Ge(001) surface reconstruction is considered, i.e. the (2 × 10) for $N = 1$, and the (2 × 8) for higher coverages; full circles: surface energy of Ge/Si(105) as a function of the number of Ge(105) layers (top axis), or of the equivalent number of Ge(001) layers (bottom axis). (b) Surface energy as a function of biaxial strain for Ge(001) in two different reconstructions, and for Ge(105)–(2 × 1). From [46], reprinted with the authors' permission. Copyright (2005) by the American Physical Society.

surfaces. The latter was investigated by several authors. It was shown that the peculiar rebonded-step (RS) reconstruction, present on the facets of actual (105) pyramids on Si(001) (see Fig. 10.7, taken from [42]), induces a strong lowering in γ with strain [45, 46, 55, 61], to the point that at the Ge/Si misfit spontaneous (105) faceting of a (001)WL becomes nearly favored (the thermodynamic preference for one of the two orientations is hard to establish based on extremely close values – see Fig. 10.6 and the discussion contained in [47]). The above calculations clarified why {105} facets seem to appear so frequently as soon as Ge is deposited on Si, not necessarily on planar Si(001) [43, 44]. The other interesting result brought up by DFT calculations is that, provided that the most energetically favorable surface reconstruction takes place, γ values seem to depend weakly on the orientation.



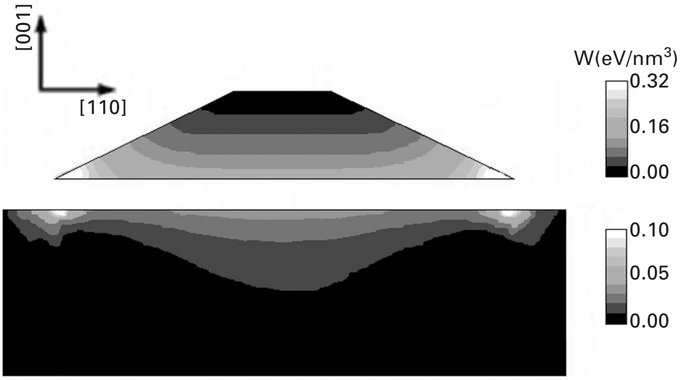
10.7 (a) Top view of the {105} surface as cut (left) and RS-reconstructed (right). Filled or empty circles are used for atoms on different terraces. To build the RS reconstruction (whose unit cell is highlighted by a dashed rectangle), encircled atoms must be removed. (b) STM images taken for both filled and empty states along the facets of {105} pyramids, compared with DFT-simulated images [42]. From [42], reprinted with the authors' permission. Copyright (2005) by the American Physical Society.

Beside the above-commented similarity of (001) and (105), also Ge(113) and Ge(111) seem to be characterized by similar values (around 62–63 meV/Å^2, see [63]), although the dependence on strain was never analyzed.

If reliable evaluation of surface energies requires using DFT, strain relaxation can be handled with less demanding approaches. Luckily so, since evaluating ρ_{is} in Eq. 10.3 by DFT for a complex multifaceted island such as a dome goes far beyond present computational possibility (clearly, instead, the WL ρ_{t} term can be directly computed by DFT due to the very simple geometry). Molecular dynamics simulations based on empirical potentials have been used to compare the deformation field associated with Ge islands (see, e.g., [32]), even when capped with Si [64]. While nowadays it is possible to handle millions of atoms on a nice desktop-PC, MD calculations are still rather demanding, so that continuum elasticity theory calculations solved by finite element methods (FEM) are often used instead. Beside the huge gain in computational speed allowed for by the use of FEM, another advantage with respect to MD is given by the possibility to directly use experimental, or DFT-derived elastic constants. On the other hand, errors could arise if linear elasticity theory is used [50], or if islands are so small that atomic-scale effects, i.e. due to surface reconstructions, are relevant. In our experience, the latter are negligible whenever islands with bases exceeding ~10–15 nm are considered (experimentally observed islands are typically larger – see Fig. 10.1). FEM allows one to quickly investigate trends. The progressive strain relaxation with increasing aspect ratio, for example, is reviewed and discussed in [28]. As the in-plane projection of the facet normal increases, Ge atoms are able to relax their effective lattice parameter in the growth plane, going beyond the simple tetragonal deformation characterizing the WL. In doing so, the island bends and induces a deformation also in the substrate. A typical distribution of elastic energy within a dome-shaped island and in the substrate below is shown in Fig. 10.8.

Whatever method is used to establish the volumetric elastic energy density ρ_{is} in Eq. 10.3, it is important to stress that ρ_{is} must include also the above-described (see also Fig. 10.2) elastic deformation experienced by the substrate. In other words, $\rho_{\text{is}}V$ must account for the total elastic energy introduced in the system due to the presence of the island. Calculations should in principle be repeated for each different volume, since the assigned WL thickness breaks self-similarity. For islands whose apex is sufficiently high (typical experimental islands fall in this category) with respect to the WL free surface, however, the deformation field in the substrate extends far beyond the WL thickness and calculations can safely be performed by placing the island directly on a Si substrate, restoring self-similarity.

Once the various parameters entering Eq. 10.3 are known, the island stability can be established as a function of N and V, for different island shapes. Truly, some additional effort is needed, since in the above discussion we have not



10.8 Elastic-energy distribution in a dome-shaped island on a Si substrate as obtained by FEM calculations. In the plot, the island is artificially lifted from the substrate in order to make clear that two different gray scales are used for the island and for the substrate. In the calculation, the island/substrate misfit was set to 0.012, therefore modeling a $Ge_{0.3}Si_{0.7}$ island, with uniform Ge distribution. From [65], reprinted with the authors' permission. Copyright (2009) by the American Physical Society.

mentioned possible problems arising from mixing parameters established with different techniques. The quantity $\rho_{\mathrm{eff}}\ (N + 1)$ is determined based on DFT-derived results, while ρ_{is} is computed by FEM calculations (here we follow [50]). Finally, ρ_{t} can be computed by using either approach. The quantitative comparison of the three volumetric energy densities is fundamental in order to establish whether or not an island is stable at a given volume and thickness. In order to obtain meaningful results, it is important to make sure that FEM and DFT are consistent. To this goal, in [50] FEM calculations were run using DFT-derived elastic constants, and a procedure to correct for non-linear terms in the elastic energy was designed, based on imposing the condition that FEM and DFT yielded identical ρ_{t}. Consistency between FEM and DFT is needed also if one is willing to include in Eq. 10.3 strain-dependent surface energies, since this requires computing the strain field along the facets by FEM, and using DFT strain-dependent γ values [47, 50, 61].

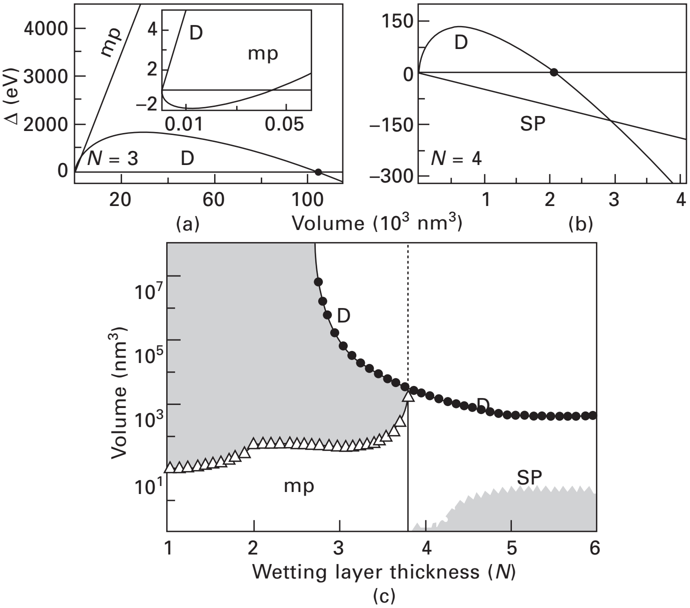
Let us now look at quantitative results for Ge/Si(001) obtained using Eq. 10.3, with DFT-derived surface energies [46] when available, FEM-derived volumetric terms, and imposing the above described DFT/FEM consistency. More details can be found in Ref. [50].

10.3.3 Unexpected theoretical predictions vs. experiments

A stability diagram for {105} pyramids and domes (whose realistic geometry was taken from Ref. [24]) vs. WL thickness and island volume is reported in

Fig. 10.9. In Figs 10.9(a) and (b) Δ, computed based on Eq. 10.3, is plotted as a function of the island volume. The D-curve represents domes (since the surface energy of some of the dome facets is unavailable from DFT, the value $\gamma_{\mathrm{dome}} = 65$ meV/Å^2 was used, verifying that reasonable changes did not alter the results significantly [50]).

For both $N = 3$ and $N = 4$, the dome curve follows the typical nucleation behavior already discussed in Section 10.3.1 (see Fig. 10.5(a)). Volumetric terms favor dome formation against WL thickening, surface terms acting oppositely. Starting from the critical volume indicated by a full circle, the energy of the dome is lower than the WL one. Such volume is reported in the D-curve of Fig. 10.9(c) as a function of the WL thickness, non-integer N values being obtained by linear interpolation. Below ~2.7 ML, there is no finite critical volume since the volumetric term becomes positive (leading to $\Delta(V)$ curves similar to the one displayed in Fig. 10.5(c)) owing to the energetic advantage provided by making the WL thicker, therefore exploiting the still strong lowering of γ_{WL} (N) (see Fig. 10.6). Notice the sizeable decrease



10.9 (a) $\Delta(V)$ as given by Eq. 10.3 for domes (D curve) and pyramids, which are shown to be metastable (mp). The WL thickness is fixed at $N = 3$. (b) As (a) but for $N = 4$. Pyramids are now stable islands (SP). (c) Critical volume vs. N for domes (D) and pyramids. In the latter case, islands are either stable above (SP) or below (mp) the critical volume. In the gray region, neither domes nor pyramids are stable. From [50], reprinted with the authors' permission. Copyright (2009) by the American Physical Society.

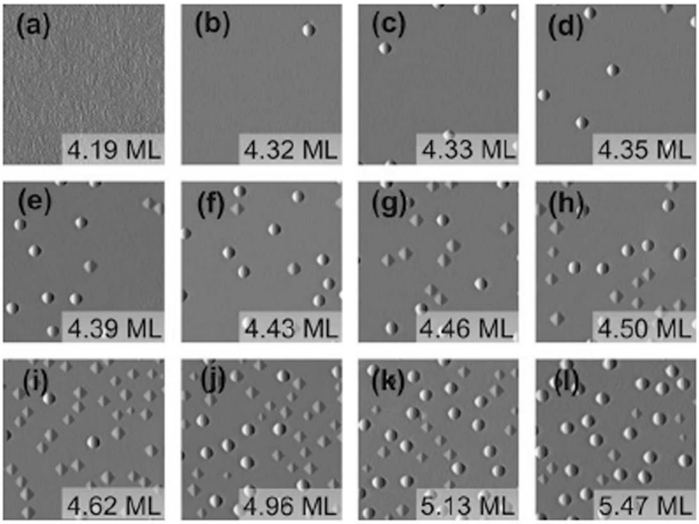
in critical volume with increasing N, until the thick film limit is reached. Pyramid behavior is more complex. From Fig. 10.9(a), it is clear that at $N = 3$ (and lower, not shown) the $\Delta(V)$ curve fully resembles the one sketched in Fig. 10.5(d). If at $N = 3$ the strain relief provided by the steep geometry of the domes was sufficient to make the volumetric term negative in Eq. 10.3, this is not true for pyramids. On the other hand, the {105} surface energy is lower than the WL one (though still rather high at low coverages – see Fig. 10.6), so that the surface term drives island formation. As a result, the $\Delta(V)$ curve has a minimum: at this coverage pyramids are stable only up to a maximum size, the opposite behavior to that of domes. While here we are discussing stability only from a thermodynamic point of view, it is easy to envision a short lifetime for these small pyramids, destabilized by any attempt to increase their volume [50]. As a consequence, we shall call pyramids in this status 'metastable pyramids' ('mp' in Fig. 10.9, as opposed to 'SP' used to indicate stable pyramids). For $N = 4$ also for pyramids the volumetric term becomes negative. Due to their low surface energy, however, their nucleation appears barrierless, the $\Delta(V)$ curve becoming the analog of Fig. 10.5(b). In a very narrow range of coverages, all four possible behaviors indicated in Fig. 10.5 are found in the Ge/Si(001) system.

Let us now consider both pyramids and domes in Fig. 10.9(c), where critical volumes (for mp these are the volumes below which pyramids are more stable than the WL) are reported as a function of N. For small coverages no stable islands exist, only mp. The first thermodynamically stable islands are domes, which can form starting from $N \sim 2.7$ ML. However, at this critical thickness, their nucleation is impossible since an infinite number of atoms would need to be gathered in a single island. As the coverage grows, dome formation becomes increasingly simple. Also, the volume of mp is predicted to increase with N. This suggests mp to be possible precursors of domes, providing an intermediate shape from which the volumetric gap leading to stable dome formation can be closed also prior to stable-pyramid formation, which occurs from $N \sim 4$ ML. From this coverage pyramids are predicted to be the first (in terms of volume) stable islands, the dome shape becoming the lowest-energy one only at sufficiently large volumes. This is the situation expected in the thick-film limit (Fig. 10.5(a)) and seems to nicely explain the typical experimental evidence. However, Fig. 10.9 shows a much greater complexity which can be tested only against very carefully designed experiments.

In [50], MBE growth of Ge on Si(001) was analyzed with a resolution of 0.025 ML. This was achieved by turning off substrate rotation during growth (an analogous procedure was used in [66] to investigate InGaAs/GaAs growth), leading to a gentle variation of Ge coverage across the wafer, which was then sectioned and analyzed. The combined use of AFM and PL allowed for a very fine control of both island shapes and WL thickness (see below). If not otherwise specified, the growth rate was set to $R = 0.05$ Å/s. In Fig. 10.10, a sequence of AFM images recorded at different coverages for a growth temperature $T = 700°\text{C}$ is reported.

Up to $N \sim 4.2$ ML no islands are detected, while at around $N \sim 4.3$ ML only domes are seen. Finally, at higher coverages, also pyramids appear. Nucleation of domes prior to pyramids seems in excellent agreement with the model prediction, under the assumption that small mp are quickly destabilized by kinetics, due to the high temperature. One could even speculate that the later appearance of pyramids is also in agreement with Fig. 10.9(c). However, one must consider that right after dome formation deep trenches are excavated leading to strong Si/Ge mixing [50], so that the system enters a condition outside the model's predictive power.

Very interestingly, experiments were repeated at lower temperatures. At $T = 625°\text{C}$ small pyramids were observed between 2 and 4.6 ML, their dimensions only slightly increasing in this rather large coverage range [50]. Only after dome formation were much larger pyramids detected. These observations also seem to reinforce our theoretical interpretation. At sufficiently low temperatures, mp have a sufficient lifetime to be experimentally observed. Also in agreement with the experiments, their volume is predicted to increase but only slightly, until they become stable and their enlargement is limited only by kinetics. Further experiments were designed to fully demonstrate that domes are the first stable islands to appear, small pyramids observed

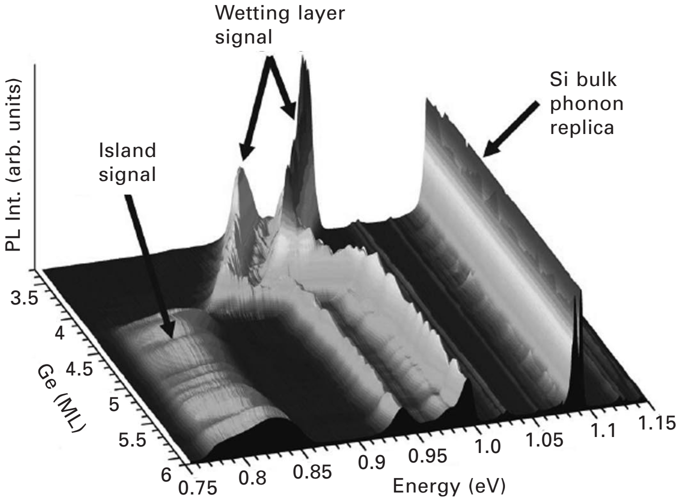


10.10 Sequence of AFM images reporting island formation and evolution as a function of the WL thickness. Illustrated samples were grown by MBE at 700°C. From [50], reprinted with the authors' permission. Copyright (2009) by the American Physical Society.

at low temperature being simply metastable. To this goal, several samples grown at $T = 625°C$ were annealed for 160 min at $T = 700°C$ in order to limit kinetic effects. Pyramids nucleating between 2 and 3.2 ML simply disappeared (no stable islands exist at this low coverage), while the ones nucleating above 3.2 ML generated domes. This evidence fully confirms the model prediction, at least from a semi-quantitative point of view.

Also, by repeating the experiments at $T = 700°C$ with the deposition flux being reduced to $R = 0.01$ Å/s, domes appeared at a reduced (by ~0.3 ML) critical thickness. The latter observation is easily explained using the model. According to Fig. 10.9(c) nucleation of stable domes could occur already around 2.7 ML, but kinetic limitations (mainly in terms of effective capture area) necessarily drive nucleation at overcritical thickness, i.e. at the $N >$ 2.7 ML value of the curve (not determined by the model) where the critical volume becomes lower than the one allowed by kinetics. Higher temperatures or lower deposition rates reduce kinetic limitations and enlarge the effective capture areas, making it possible to nucleate stable islands at thicknesses close to the minimum one allowed for by thermodynamics.

While sequences of AFM images such as the one displayed in Fig. 10.10 offer a direct idea of the system evolution with deposition, they do not allow the WL behavior to be monitored during growth. To this goal, PL measurements have been conducted, leading to the results shown in Fig. 10.11. Simultaneous inspection of Figs 10.10 and 10.11 therefore offers a



10.11 PL spectra recorded at different coverages, and grouped together in a three-dimensional plot. The Ge growth temperature was the same as that used to produce the results of Fig. 10.10. From [50], reprinted with the authors' permission. Copyright (2009) by the American Physical Society.

clear experimental picture of the system evolution at this temperature. When no islands are revealed by AFM, one observes in the PL spectra only the WL signal, the shift toward lower energies signaling lesser confinement and, thus, WL thickening. At around 4.2 ML, i.e. exactly when domes are imaged by AFM in Fig. 10.10, two main features distinguish the PL signal. First of all, the islands (domes) signal sets in. Secondly, an abrupt blueshift is clearly seen. Using both this experimental indication and the theoretical results of Fig. 10.9, the interpretation becomes rather straightforward. We have already discussed why islands are expected to nucleate under some overcritical-thickness conditions, due to kinetic limitations. Once formed, however, they can acquire material from the WL which is therefore considerably thinned (by almost 1 ML in Fig. 10.11). This process is consistent with thermodynamics since it corresponds to Fig. 10.9(c) in following the dome curve from right to left. A quantitative estimate of the actual amount of material transferred from the overcritical WL to the islands would need an explicit treatment of kinetics or, at least, the assignment of typical WL capture areas to each individual island. While this is left for future work, we notice that Fig. 10.9(c) directly supplies an upper limit for the WL thinning, in showing that the process must stop when the WL reaches the minimum (i.e. ~2.7 ML) critical thickness for dome formation.

The complex behavior illustrated above for Ge/Si(001), where islands nucleate at overcritical WL thicknesses and subsequently enlarge at the expense of the WL, has also been observed in InGaAs/GaAs(001) [66], and theoretically investigated using a conceptually very similar approach (a combination of DFT surface energies with FEM elastic terms) [67, 68].

10.3.4 Model limitations

By looking at the dome curve displayed in Fig. 10.9(b), one sees that the simple thermodynamic model predicts indefinite growth of such islands, once the critical volume for their nucleation is reached. This is obviously an artifact of the model which is designed solely to describe the very early stages of growth, i.e. the onset of SK island formation. In a real system a number of different phenomena oppose dome growth beyond a certain volume. First of all, in Fig. 10.9 only two island shapes have been considered, but transition from domes to steeper morphologies such as barns [18] would be predicted at sufficiently large volumes, owing to the larger volumetric term. Furthermore, always at sufficiently large volumes [24, 25], plastic relaxation would become competitive. Also, the model does not consider the lateral repulsion between islands [8, 48], which surely plays a role at advanced deposition stages, when the island density is large. Under these conditions, also coarsening and ripening should be considered [35].

It is clear that only a simulation including kinetics and providing the actual time evolution of the system during growth could capture the whole complexity of SK growth. Atomistic approaches in terms of kinetic Monte Carlo (KMC) simulations have been attempted by several authors. Although pioneering work was presented by Madhukar as early as 1983 [69], more recents developments can be found, e.g., in Refs [70–72]. Performing a KMC simulation of Ge/Si growth at experimental time and length scales is a formidable task, considering the need for inserting the long-range elastic field. Unavoidably, approximations need to be introduced, for example in terms of geometry. Typically, the popular use of solid-on-solid models, while allowing a considerable quickening of the simulation, does not allow the capture of details such as surface reconstructions which are important, as previously demonstrated, to predict the correct island shapes and their stability. As a consequence, KMC simulations to date have revealed many qualitative features of SK growth, but considerable work is still required before reaching the predictive power needed, for example, to explain the complex evolution reported in Figs 10.10 and 10.11 Longer time and length scales can be reached by using continuum models, where the key driving force for material motion via surface diffusion is described in terms of surface gradients of μ, the chemical potential calculated at the surface [73]. To correctly treat heteroepitaxial growth, μ must include strain effects and curvature-driven diffusion. Besides, to predict islands with realistic shapes, a critical orientation-dependent surface energy term must also be present. Nice examples of continuum simulations revealing various aspects of SK growth can be found in Refs [74, 75].

In the above-quoted simulations of SK growth, as well as in our simple thermodynamic model, a further important ingredient is missing, i.e. time-dependent Si/Ge intermixing. Dome islands shown in Fig. 10.10 are heavily intermixed, the average Ge concentration c being estimated to be ~0.35 [50]. As already stressed, most of the Si entering the islands comes from the trenches around their periphery, after islands become sufficiently steep (i.e. domes) [38]. The theoretical prediction of Fig. 10.9 can therefore be considered as reliable only as soon as domes are first formed. In the experiments, trench excavation and Si enrichment are so fast that once domes can be imaged, the process has already taken place. This makes the comparison between theory and experiments only qualitative in terms of volumes. Intermixing, indeed, causes a significant shift of critical volumes towards larger values. This can be immediately understood using Eq. 10.2. If variations in surface energy with concentration are neglected, the condition $\Delta(V) = 0$ implies $V \approx D^{-3}$, and thus $V \approx c^{-6}$. Extending the thermodynamic model to consider variations in concentration is possible in principle, but would unavoidably lead to insertion of adjustable parameters, ruining the attempt to predict the Ge/Si(001) behavior fully based on known, theoretically calculated quantities. Notice that beside pure-Ge islands, the model also considers a pure-Ge WL.

Still, the latter assumption appears less severe, owing to the high Ge content revealed also at high growth T [50, 59]).

10.4 Beyond the Stranski–Krastranow (SK) onset: SiGe intermixing

Capturing the complexity of Ge–Si intermixing during growth is one of the current main goals of the theoretical community. The task is difficult, since alloyed islands do not present homogeneous distributions of Ge in their interior. This was nicely demonstrated by a large set of recent experiments, where attempts to determine the 3D distributions within SiGe islands were presented, based either on X-rays (for recent examples see [76–78]) or on selective chemical etching [79–81].

Direct simulations of the system time evolution during growth, including intermixing, are rare in the literature. From the atomistic point of view, a method to include alloying in a convenient way was presented only very recently [82]. Nice continuum simulations which seemed to capture most of the actual complexity, notably considering that system equilibration in terms of Ge/Si exchanges is possible only in surface and subsurface layers (again, the local thermodynamic concept), were presented in [35]. This model appears very promising although still restricted to 2D and based on the use of the flat-island approximation.

More realistic results, in terms of description of island shapes and elastic fields, can be obtained if instead of solving the system dynamics, one is willing to answer a simpler question, i.e. what is the Ge–Si distribution minimizing the free energy G, for an assigned shape and average $c \equiv \bar{c}$. There are three main contributions to G to be considered: the elastic energy E_{el}, the entropic term $TS = -k_B T\,[c \log c + (1 - c) \log (1 - c)]$ (here and in other works [77, 83, 84] taken for simplicity as the entropy of mixing of an ideal mixture), and the Ge/Si enthalpy of mixing, H [85]. The role of S and H is clear, and one is opposed to the other. Entropy maximization forces uniform Ge distributions, while H induces full segregation. Since H is rather small for Ge/Si [85], at typical experimental temperatures the combined effect of mixing entropy and enthalpy (to a first approximation, the latter simply provides an effective lowering of temperature, by some 300 K [85], in the TS term) is to provide constant c across the island: $c(x, y, z) = \bar{c}$. The role of E_{el} is more complex, and can be understood by looking for the Ge distribution $c_{\min}(x, y, z)$ minimizing the elastic energy of the system. This (and, more generally, free-energy minimization) can be accomplished by atomistic Monte Carlo simulations based on empirical potentials, where an initial Ge distribution is assigned, pairs of Si/Ge atoms are exchanged, and moves are accepted based on the statistical Boltzmann factor [86, 87]. The computational effort is, however, considerable, preventing one from systematically investigating trends in terms of different island shapes and $\bar{c}$ values.

Continuum approaches have therefore been designed. In [84, 88], the MC approach is kept, but within a FEM, continuum-elasticity scheme, providing a significant computational gain with respect to the atomistic approach. Even faster is the recent method of Ref. [65] (but see also [89] for an alternative approach), which we briefly outline below. The starting point is the general equilibrium equation for an elastic body (simply stating that the sum of all internal and external forces should be zero, taking into account proper boundary conditions – see [90]):

$$\begin{cases} -\sigma_{ij,j} = f_i & \text{on } \Omega \\ u_i = 0 & \text{on } \Gamma_\mathrm{D} \\ \sigma_{ij} n_j = g_i & \text{on } \Gamma_\mathrm{N} \end{cases} \tag{10.4}$$

where $i(j) = 1, 2, 3$; u_i is the ith component of the displacement field; σ (which depends on the displacement field) is the stress tensor; f_i represents a volumetric force density applied to the body and g_i a surface force; Ω is the domain occupied by the elastic body; and $\Gamma_D \cup \Gamma_N$ is its total boundary within which Neumann boundary conditions (BC) are imposed on Γ_N and Dirichlet BC on Γ_D. Furthermore, n_j is the jth component of the local surface normal to Γ_N, the symbol $(_{,j})$ representing derivation with respect to the jth coordinate, and summation of repeated indexes is assumed.

When treating lattice-mismatched islands on a substrate, it is particularly convenient to use Eshelby's formalism. This means that the island is seen as an Eshelby inclusion [91], and the solution of Eq. 10.4 is found by solving the homogeneous problem (null volumetric and surface forces), using the modified stress–strain relation $\sigma_{ij} = C_{ijkl}(\varepsilon_{kl} - \varepsilon^*_{kl})$ where ε is the elastic strain tensor, and ε^* the eigenstrain mimicking the island/substrate misfit f, $\varepsilon^*_{ij} = f\,c(x, y, z)\delta_{ij}$. The elastic energy reads:

$$E_\mathrm{el} = \frac{1}{2}\int_\Omega \mathrm{d}V\,(\varepsilon_{ij} - \varepsilon^*_{ij})C_{ijkl}(\varepsilon_{ij} - \varepsilon^*_{ij}) \tag{10.5}$$

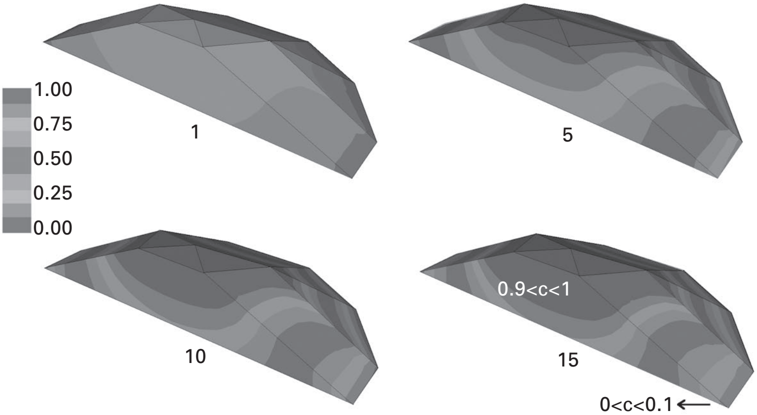
From Eq. 10.5 one sees that E_el is a functional of ε and ε^*, both depending on $c(x, y, z)$. Minimization with respect to the eigenstrain for fixed ε can be tackled analytically. Starting from an initial distribution c, it can be shown [65] that the variation

$$\Delta c = \frac{1}{3f}(\mathrm{Tr}\,\varepsilon - \overline{\mathrm{Tr}\,\varepsilon}) + (\bar{c} - c) \tag{10.6}$$

where $\overline{\mathrm{Tr}\,\varepsilon}$ is the trace of the strain tensor averaged over the elastic body, provides $E^\varepsilon_\mathrm{min}$, i.e. the minimum elastic energy for fixed elastic strain ε.

Clearly, $E^{\varepsilon}_{\min}$ is not the true minimum elastic energy of the system, since the elastic response to a change in the Ge distribution (i.e. to a variation in the eigenstrain) must be added. This step requires solving the elastic problem, e.g. by using a FEM solver, a computationally expensive step. Luckily, there is a procedure whereby, starting from an initial Ge distribution $c_0(x, y, z)$, Eq. 10.6 is used to predict a new $c_1(x, y, z)$, the FEM solver is called to update the elastic strain ε and Eq. 10.6 is exploited again, leading to rapid convergence to the actual $c_{\min}(x, y, z)$ minimizing the elastic energy, as illustrated in Fig. 10.12 for a dome-shaped island with $\bar{c} = 0.5$. From the qualitative point of view, interpretation of the converged profile reported in Fig. 10.12 is straightforward. Ge is preferentially accumulated at the top of the island, where the lattice parameter is closer to Ge-bulk. Conversely, the compressive strain is produced at the bottom edges of an island with uniform distribution (see the high-energy region in Fig. 10.8).

Ge profiles similar to the one displayed in Fig. 10.12 were reported also in Ref. [88] and shown to closely resemble the experimental etching profiles of Ref. [80], obtained at low growth temperatures (580°C). Interpreting this agreement is not simple. First of all, if minimization is carried out for $\bar{c} < 0.4$, i.e. at typical average values characterizing islands grown at 700°C or above, the agreement between theory and experiments is lost [50, 81], elastic-energy minimization predicting too high Ge contents at the top of the island [88]. This could lead to thinking that the main problem resides in minimizing elastic energy instead of the full free-energy functional. Inserting



10.12 Evolution of the Ge concentration profile during the iterative procedure leading to $c_{\min}(x, y, z)$ for a dome-shaped island with an average Ge content $\bar{c} = 0.5$. The number of FEM calls is reported below each of the panels. After 15 FEM calls, convergence is reached. On the left the scale for the color map is reported: 1.00 means 100% Ge. From [65], reprinted with the authors' permission. Copyright (2009) by the American Physical Society.

entropy, however, one finds Ge distributions almost fully uniform at the typical growth temperatures, including the previously discussed $T = 580°\text{C}$ [84] case.

The main problem stems once again from the care needed when using thermodynamic arguments. In determining optimal concentration profiles, we (and others, see e.g. [86, 89]) tacitly assumed that atoms within the islands were able to migrate to the most energetically convenient position in the island, something which appears impossible given the absence of bulk diffusion [3]. Comparing predicted thermodynamic profiles with experimental ones is, however, extremely interesting, since it allows one to understand how kinetic limitations act. The full set of results described above points out that kinetics limit entropy maximization more than elastic energy minimization, as if an effective temperature lower than the experimental one should be used when minimizing the free energy. Recent DFT calculations may offer some hints in this respect. In Ref. [36], indeed, atomic Si/Ge exchanges at {105} facets were analyzed. The kinetic barrier to be surmounted by a Si adatom for incorporating in a subsurface stable site of the facet (creating a Ge adatom) turned out to be comparable with the diffusion one [1, 92]. This means that if a small, pure-Ge island is created, as soon as trenches are formed around it and Si atoms start climbing over the facets, their diffusion toward the island top is slowed down by an incorporation process, yielding a purely kinetic justification for the observed profiles. Clearly, this is only an indication of the actual behavior, since calculations are limited to the {105} orientation. More work is surely needed in order to fully clarify this issue and others not discussed in this section. For example, we notice that the 'age' of an island could play a role in determining its composition, since the first ones to form find a Ge-richer environment due to the absence of trenches. Only a suitable kinetic model could explain this further complexity.

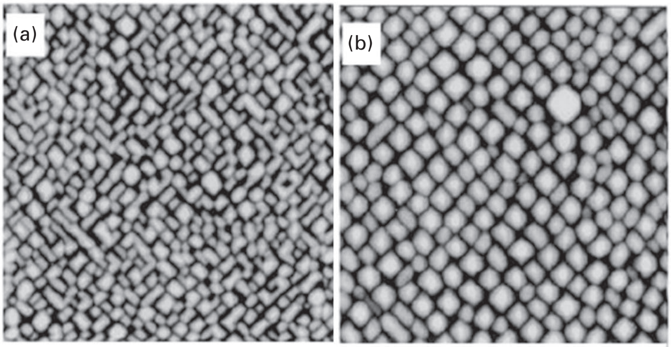
10.5 Beyond the Stranski–Krastanow (SK) onset: vertical and horizontal ordering for applications

10.5.1 Background and the Tersoff, Teichert and Lagally (TTL) model

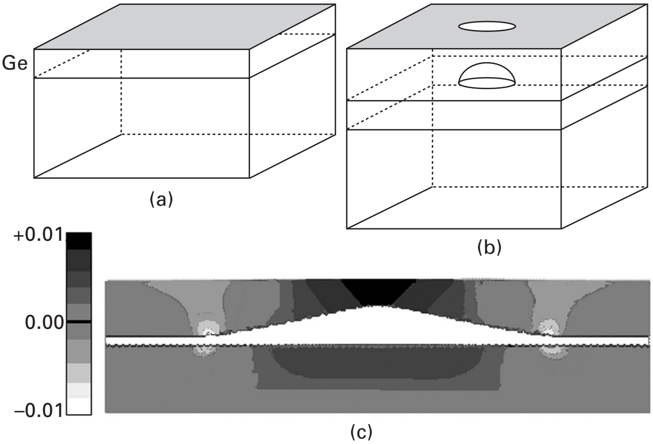
In order to exploit islands for applications, it is mandatory to control their shape, size, composition, and positioning. When growing on Si(001), the last requirement appears particularly difficult to achieve. Indeed, on an ideally flat Si(001), islands will nucleate randomly along the WL, due to the initial uniformity of the strain field. If one is not willing to pattern the substrate, and aims at obtaining ordering by pure self-assembly, creation of stacked multilayered structures provides a solution. This was demonstrated by Tersoff, Teichert, and Lagally (TTL) in [93], where positioning of a single layer of islands obtained by depositing $Si_{0.25}Ge_{0.75}$ (Fig. 10.13(a)) on Si(001) was compared with a superlattice where each layer of islands was capped with a pure-Si spacing layer (SL), before growing the alloy again (Fig. 10.13(b)). The enhanced degree of ordering is evident.

The interpretation of Fig. 10.13 is not straightforward. The presence of buried SiGe islands induces a strain modulation at the surface of the Si SL. Since the upper region of an island is more relaxed (see Fig. 10.8), the portion of the Si SL closer to the island top will be expanded, providing a preferential site for further island nucleation. The situation is sketched in Fig. 10.14, taken from [95], where an actual MD calculation of the strain field of a buried island is also included. What Fig. 10.14 directly suggests is vertical ordering of islands: the strain field clearly drives piling up of islands in different layers. Cross-sectional TEM images are needed to see the vertical arrangement of islands in a multilayered structure. While not present in [93], TEM evidence of piling up was supplied later for similar samples (see [96] and references therein, and [97] where vertical ordering for III/V semiconductors was first seen). While we have explained why growth on a layer of buried islands and on a flat substrate are different, we still need to understand why islands appear ordered in the growth plane in Fig. 10.13. A simple one-dimensional model for the process was provided in [93], where the strain field induced by a buried island at the SL surface was modeled analytically. This is possible provided that islands are treated as force dipoles, so that within classical elasticity theory [98] one finds the simple expression [99]:

$$\mathrm{Tr}\,\varepsilon = CV(r^2 + H^2)^{-3/2}\left(1 - \frac{3H^2}{r^2 + H^2}\right) \qquad 10.7$$



10.13 (a) AFM images of a single layer of islands grown by depositing $Si_{0.25}Ge_{0.75}$ on Si(001). (b) Upper layers of islands after depositing 20 layers of islands separated by a Si spacer. From [93], reprinted with the authors' permission. Copyright (1996) by the American Physical Society.

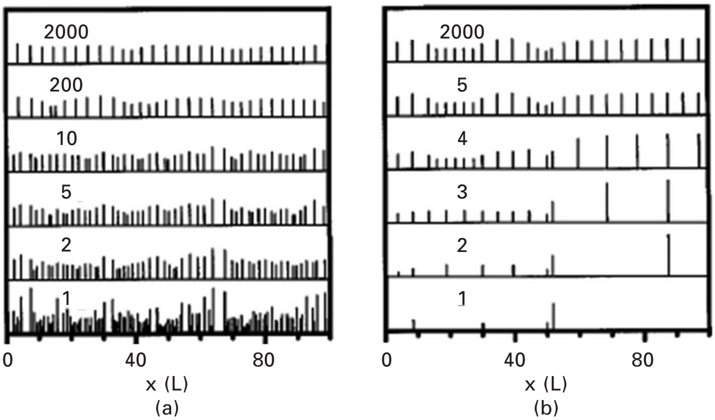


10.14 On a flat WL (a) island nucleation is random (any point of the gray region is equivalent). On top of Si-capped islands (b), instead, a preferential nucleation region is present (white area within the gray region). In (c) the ε_{xx} component (x being the horizontal direction in the panel) of the strain tensor, computed using MD simulations based on the Tersoff potential [94], is plotted for a Ge {105} pyramid on a WL, capped with Si. The scale is chosen to emphasize the deformation at the surface of the capping layer. As a consequence, details of the strain behavior within the buried islands are not visible. Panels (a) and (b) are reprinted from [95] with the authors' permission, copyright (2007) by the Institute of Physics. Panel (c) is reprinted from [64] with the authors' permission, copyright (2005) by the American Institute of Physics.

where H is the SL thickness, r is the distance from the island in the surface plane, and C depends on the elastic constants. TTL showed that superposition of the field described by Eq. 10.7 induces progressive horizontal ordering. By placing islands in points where Eq. 10.7 yields a lattice parameter closer to Ge, the evolution here reported in Fig. 10.15 is obtained: vertical ordering is accompanied by progressive horizontal ordering. If two islands are very close, indeed, the superposition of their elastic fields creates a single preferential nucleation point in the upper layer, while if two islands are very far apart, the tails of the strain field create a (weakly) favored nucleation region in between [93, 95].

10.5.2 Extending the TTL model: shape effects

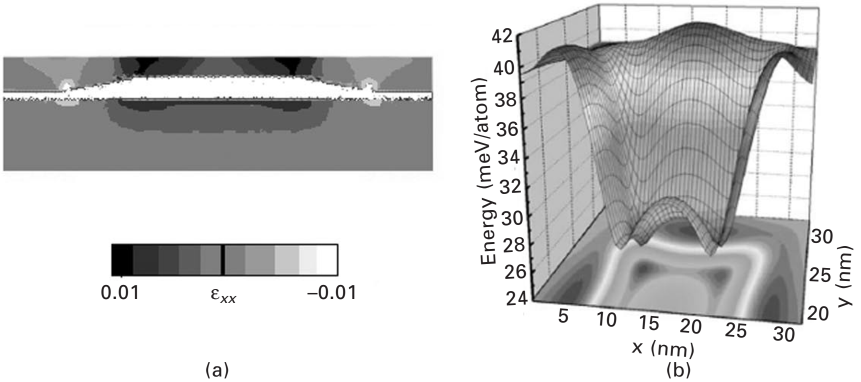
After the appearance of Ref. [93], several authors investigated vertical and horizontal island ordering. Extensive reviews can be found in Refs [8, 10, 11]. Here and in Section 10.5.3 we wish to report on some recent results which helped to extend the simple TTL model. The first interesting aspect



10.15 (a) Starting from a distribution of closely spaced, disordered islands, horizontal ordering is achieved after a few island layers. (b) Similar ordering is achieved also by starting from fewer, widely spaced islands. In both panels, the height of the vertical segments is proportional to the island volume, and the island layer is indicated. The island position *x* is measured in units of the SL spacing *L*. From [93], reprinted with the authors' permission. Copyright (1996) by the American Physical Society.

resides in relaxing the hypothesis of buried islands as point-like sources of strain. In [64] MD simulations were used to compute the strain field induced by a buried island of realistic shape (but see also [100]). During capping with Si, dome islands tend to revert back to their pyramidal shape [101], down to even lower aspect ratios caused by the loss of absence of a real apex [64]. The presence of a truncated (001) termination causes a peculiar feature in the strain field at the SL surface, as demonstrated by the results displayed in Fig. 10.16 (compare with Fig. 10.14(c) where a complete pyramid is considered).

Nucleation over the energetic landscape pictured in Fig. 10.16(b) looks interesting since, provided that islands are small, nucleation right on top of the buried dot appears unlikely, due to the presence of a local maximum in the atomic chemical potential. In CVD experiments where the growth temperature was kept sufficiently low [64, 102] to reduce the volume of the upper-layer islands, indeed, clear evidence of islands closely clustered around an inner void zone was found. As expected and predicted by the theory (see Eq. 10.7 and [64]), this peculiar arrangement is lost for thick SLs. Interesting KMC simulations exploring further the possibility of exploiting the modulated strain field at the SL surface to force peculiar island positioning can be found in [103].

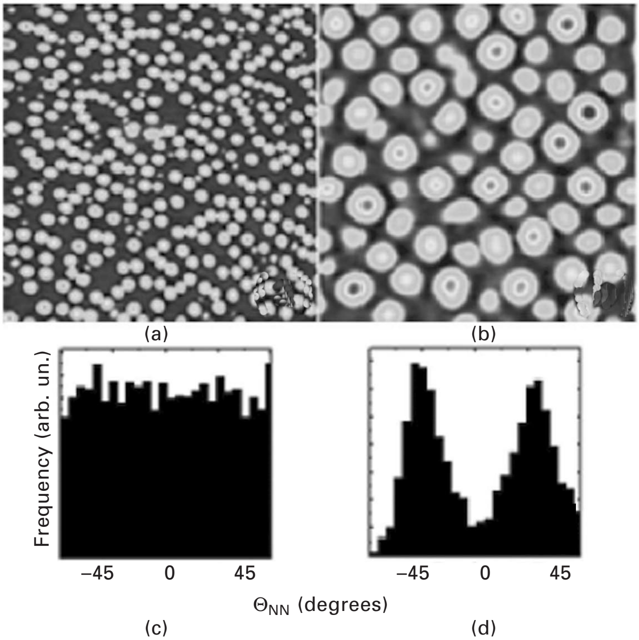


10.16 (a) Strain distribution induced in the Si SL by a truncated pyramid, computed using MD simulations based on the Tersoff potential [94]. (b) The corresponding chemical potential at the SL surface. The central maximum is located directly above the (001) top facet of the truncated island. The SL thickness was set to ~1/10 of the pyramid base. The maximum disappears for a SL twice as thick [64]. From [64], reprinted with the authors' permission. Copyright (2005) by the American Institute of Physics.

10.5.3 Extending the TTL model: ordering from the first layer induced by moderate capping

In the TTL model horizontal ordering is achieved progressively layer by layer as a result of the strain field produced by buried islands at the SL surface. Recent results indicate that ordering can take place also during the capping process. In [104], indeed, a first layer of islands was created by CVD, revealing the usual random distribution of islands. Partial covering by only a few nanometers of Si was then demonstrated to produce a significant rearrangement of the islands, displaying a degree of ordering dependent on the growth temperature (for fixed capping temperature). In the example reported in Fig. 10.17, the first layer of island was created upon depositing GeH_4 at the relatively low growth temperature $T = 600°C$. Further deposition of SiH_4 at 750°C produced the change in island distribution demonstrated by the AFM images of Figs 10.17(a) and (b), better quantified in Figs 10.17(c) and (d) in terms of angular distribution.

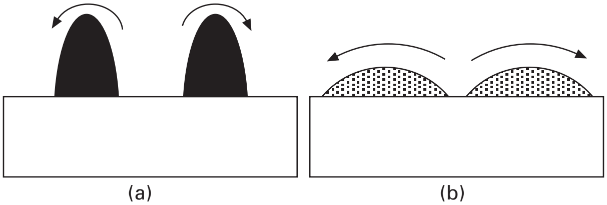
The thermodynamic driving force for the ordering process reported in Fig. 10.17 is originated by the lateral repulsive interaction between islands. As shown in Fig. 10.8, islands induce a deformation field extending in the substrate and, laterally, well beyond the island boundaries. In case two islands are closely located, superposition of two such fields raises the elastic energy of the system, acting as an effective repulsion. As discussed in [104], such repulsion destabilizes the regions of the two islands facing each other. Atoms



10.17 AFM image of the island distribution (a) before and (b) after partial Si-capping. (c) After individuating the nearest neighbors (NN) of each island, the angular distribution Θ_{NN} is built [104] and displayed (c) before and (d) after capping. Perfect arrangement in a square lattice would display two delta peaks at –45 and +45 degrees. From [104], reprinted with the authors' permission. Copyright (2006) by the American Physical Society.

will therefore migrate to the opposite side, increasing the distance between the islands and lowering the elastic repulsion. In the presence of several islands, it is easily seen how the combination of pairwise elastic repulsions can lead to ordering phenomena such as the one imaged in Fig. 10.17. Although the effect of island–island repulsion had already been discussed in [48], where islands were left uncapped, the results of [104] show that pronounced evolution towards ordered arrays needs deposition of Si. The reason for this is not trivial and, at first, the role played by Si could appear to be counterintuitive. Deposition of Si leads to flattening of the islands (islands in Fig. 10.17(a) are domes, while in Fig. 10.17(b) shallow truncated pyramids are present), caused by strong Si/Ge intermixing. Both flattening and intermixing, *per se*, reduce the deformation field.

However, flattening decreases the distance between two adjacent island edges. As demonstrated in [104] using elasticity theory, the overall effect is to destabilize more strongly the island regions facing each other. A schematic representation of the process is reported in Fig. 10.18. Notice that ordering

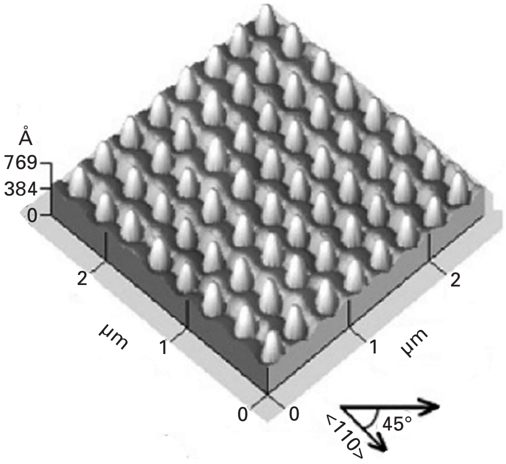


10.18 (a) Two steep islands are rich in Ge, but far apart. Elastic repulsion moderately drives migration of atoms from the facets facing each other, to the external ones. (b) Despite being shallower and less Ge-rich, the edges of the two islands are now much closer. As a consequence, the driving force for atomic migration from inner to outer regions can be stronger [104].

requires effective lateral displacements of large 3D islands. While this could appear difficult due to the large number of atoms involved (of the order of 10^4–10^5 or more), demonstration of island motion was given in [105]. Using selective chemical etching, the authors analyzed the deformation left on the substrate by large (around 200 nm in base) dome islands upon extended annealing. In some cases islands were observed to move despite being rather isolated. An explanation in terms of individual-island elastic energy and entropy was given. It was shown that any fluctuation in Ge content between opposite sides of the island can drive a different chemical potential involving transfer of Ge atoms from the Ge-richer side to the other, accompanied by capture of Si atoms from the substrate. This effect leads to an effective lateral motion during which islands grow in volume and become progressively diluted, until there the driving force for motion goes to zero. This mechanism appears to be complementary to the one proposed in [104]. Due to the larger average distance between island edges (see [105]), island–island interaction is weaker. Indeed, no significant ordering is found to be induced by lateral motion in this case.

10.6 Future trends: ordering Ge islands on pit-patterned Si(001)

In this chapter we have discussed several aspects of SK growth of Ge islands on Si(001). We wish to conclude by briefly apidly mentioning why, based on recent results, growth on pit-patterned Si(001) substrates leads to several advantages, also in terms of understanding basic physics. The main reason for using pre-patterned substrates is clear. As we have seen in Section 10.5, it is difficult to obtain high degrees of lateral ordering by pure self-assembly. Recent developments in nano-lithography, however, allow one to carefully modify Si(001) substrates, creating arrays of ordered preferential nucleation sites. For example, in Fig. 10.19, Ge is deposited over pit-patterned Si(001)

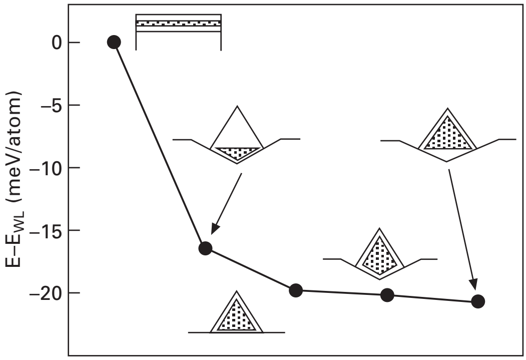


10.19 AFM 3D topography of islands grown by MBE by depositing Ge on Si(001) patterned with a 370 nm × 370 nm array of pits (growth temperature: $T = 700°C$). From [106], reprinted with the authors' permission. Copyright (2004) by the American Institute of Physics.

[106]. Beside the remarkable lateral positioning, also the uniformity in size and shape of the islands should be noticed. Later studies confirmed also a better uniformity with respect to the flat case in the Si/Ge distribution within the islands [78, 107].

Better uniformity in both shape and size can be ascribed to the different kinetics of island self-assembly in pit-patterned substrates where islands grow more independently, and under more uniform conditions [108]. Lateral repulsion, Ostwald ripening, and, more in general, neighborhood-dependent evolution can all be suppressed by a suitable patterning. With this respect, an isolated-island model such as the one presented in Section 10.3 suffers from less severe limitations if applied to a pit-patterned substrate.

Very recent results have shown that the role of pit-patterning is even more important. Preferential nucleation in pits was indeed demonstrated to be driven not only by capillarity effects (covering of the pits allows for a reduction of the exposed surface) but also by strain relaxation: a Ge island grown on a partially filled pit is significantly more relaxed [30, 78] with respect to its counterpart on flat Si(001), as illustrated in Fig. 10.20. This effect might have consequences [109] also on the process of dislocation injection. Indeed, under similar growth conditions, coherent islands in pits were observed to exist at volumes where islands are plastically relaxed on flat substrates [30]. The possibility of tuning strain relaxation and dislocation injection by a peculiar (in terms of both spacing and shape) pit patterning is surely intriguing, and likely to attract considerable attention in the next few years.



10.20 Energy per atom of the shaded regions, as obtained by MD based on Tersoff potentials [94]. Nucleation in the pit guarantees lower elastic energy. From [30], reprinted with the authors' permission. Copyright (2007) by the American Physical Society.

10.7 References

[1] F. Montalenti, D.B. Migas, F. Gamba, and L. Miglio, *Phys. Rev. B* **70**, 245315 (2004).

[2] S. Cereda and F. Montalenti, *Phys. Rev. B* **75**, 195321 (2007).

[3] B.P. Uberuaga, M. Leskovar, A.P. Smith, H. Jónsson, and M. Olmstead, *Phys. Rev. Lett.* **84**, 2441 (2000).

[4] I.T. Michely and J. Krug, *Islands Mounds and Atoms*, vol. 42 of Springer Series in Surface Science, Springer, Berlin (2004).

[5] H. Ibach and H. Lüth, *Solid State Physics: an Introduction to Principles of Materials Science*, Springer (1995).

[6] Y.-W. Mo, D.E. Savage, B.S. Swartzentruber, and M.G. Lagally, *Phys. Rev. Lett.* **65**, 1020 (1990).

[7] D.J. Eaglesham, H.-J. Gossmann, and M. Cerullo, *Phys. Rev. Lett.* **65**, 1227 (1990).

[8] V.A. Shchukin and D. Bimberg, *Rev. Mod. Phys.* **71**, 1125 (1999).

[9] B. Voigtländer, *Surf. Sci. Rep.* **43**, 127 (2001).

[10] C. Teichert, *Phys. Rep.* **365**, 335 (2002).

[11] J. Stangl, V. Holý, and G. Bauer, *Rev. Mod. Phys.* **76**, 725 (2004).

[12] I. Berbezier and A. Ronda, *Surf. Sci. Rep.* **64**, 47 (2009).

[13] A. Vailionis, B. Cho, G. Glass, P. Desjardins, D.G. Cahill, and J.E. Green, *Phys. Rev. Lett.* **85**, 3672 (2000).

[14] K.M. Chen, D.E. Jesson, S.J. Pennycook, T. Thundat, and R.J. Warmack, *Phys. Rev. B* **56**, R1700 (1997).

[15] G. Medeiros-Ribeiro, A.M. Bratkovski, T.I. Kamins, D.A.A. Ohlberg, and R.S. Williams, *Science* **279**, 353 (1998).

[16] F.M. Ross, R.M. Tromp, and M.C. Reuter, *Science* **286**, 1931 (1999).

[17] J.T. Robinson, A. Rastelli, O. Schmidt, and O.D. Dubon, *Nanotechnology* **20**, 085708 (2009).

[18] M. Stoffel, A. Rastelli, J. Tersoff, T. Merdzhanova, and O.G. Schmidt, *Phys. Rev. B* **74**, 155326 (2006).

[19] A. Rastelli, *Structural evolution of nanoscopic islands of Ge and SiGe on Si(001)*, PhD Thesis, University of Pavia, Italy (2002).

[20] F.M. Ross, J. Tersoff, and R.M. Tromp, *Phys. Rev. Lett.* **80**, 984 (1998).

[21] F.K. LeGoues, M.C. Reuter, J. Tersoff, M. Hammar, and R.M. Tromp, *Phys. Rev. Lett.* **73**, 300 (1994).

[22] T. Merdzhanova, S. Kiravittaya, A. Rastelli, M. Stoffel, U. Denker, and O.G. Schmidt, *Phys. Rev. Lett.* **96**, 226103 (2006).

[23] M.-I. Richard, T.U. Schülli, G. Renaud, E. Wintersberger, G. Chen, G. Bauer, and V. Holý, *Phys. Rev. B* **80**, 045313 (2009).

[24] A. Marzegalli, V.A. Zinovyev, F. Montalenti, A. Rastelli, M. Stoffel, T. Merdzhanova, O.G. Schmidt, and L. Miglio, *Phys. Rev. Lett.* **99**, 235505 (2007).

[25] R. Gatti, A. Marzegalli, V.A. Zinovyev, F. Montalenti, and L. Miglio, *Phys. Rev. B* **78**, 184104 (2008).

[26] C. Ratsch and A. Zangwill, *Surf. Sci.* **293**, 123 (1993).

[27] V.A. Zinovyev, G. Vastola, F. Montaleuti and L. Miglio, *Surf Sci.* **600**, 4777 (2006).

[28] G. Vastola, R. Gatti, A. Marzegalli, F. Montalenti, and L. Miglio, in *Self-Assembled Quantum Dots*, edited by Z.M. Wang, Springer, Berlin (2008).

[29] J. Tersoff and R.M. Tromp, *Phys. Rev. Lett.* **70**, 2782 (1993).

[30] Z. Zhong, W. Schwinger, F. Schäffler, G. Bauer, G. Vastola, F. Montalenti, and L. Miglio, *Phys. Rev. Lett.* **98**, 176102 (2007).

[31] G. Vastola, F. Montalenti, and L. Miglio, *J. Phys.: Condens. Matter* **20**, 454217 (2008).

[32] P. Raiteri, L. Miglio, F. Valentinotti, and M. Celino, *Appl. Phys. Lett.* **80**, 3736 (2002).

[33] D.T. Tambe and V.B. Shenoy, *Appl. Phys. Lett.* **85**, 1586 (2004).

[34] Y. Tu and J. Tersoff, *Phys. Rev. Lett.* **93**, 216101 (2004).

[35] Y. Tu and J. Tersoff, *Phys. Rev. Lett.* **98**, 096103 (2007).

[36] S. Cereda and F. Montalenti, *Phys. Rev. B* **81**, 125439 (2010).

[37] S.A. Chaparro, J. Drucker, Y. Zhang, D. Chandrasekhar, M.R. McCartney, and D.J. Smith, *Phys. Rev. Lett.* **83**, 1199 (1999).

[38] S.A. Chaparro, Y. Zhang, and J. Drucker, *Appl. Phys. Lett.* **76**, 3534 (2000).

[39] S.A. Chaparro, Y. Zhang, J. Drucker, D. Chandrasekhar, and D.J. Smith, *J. Appl. Phys.* **87**, 2245 (2000).

[40] G. Capellini, M. De Seta, and F. Evangelisti, *Appl. Phys. Lett.* **78**, 303 (2000).

[41] M. De Seta, G. Capellini, and F. Evangelisti, *J. Appl. Phys.* **92**, 614 (2002).

[42] P. Raiteri, D.B. Migas, L. Miglio, A. Rastelli, and H. von Känel, *Phys. Rev. Lett.* **88**, 256103 (2002).

[43] L. Persichetti, A. Sgarlata, M. Fanfoni, and A. Balzarotti, *Phys. Rev. Lett.* **104**, 036104 (2010).

[44] G. Chen, G. Vastola, H. Lichtenberger, D. Pachinger, G. Bauer, W. Jantsch, F. Schäffler, and L. Miglio, *Appl. Phys. Lett.* **92,** 113106 (2008).

[45] D.B. Migas, S. Cereda, F. Montalenti, and L. Miglio, *Surf. Sci.* **556**, 121 (2004).

[46] G.-H. Lu, M. Cuma, and F. Liu, *Phys. Rev. B* **72**, 125415 (2005).

[47] O.E. Shklyaev, M.J. Beck, M. Asta, M.J. Miksis, and P.W. Voorhees, *Phys. Rev. Lett.* **94**, 176102 (2005).

[48] J.A. Floro, G.A. Lucadamo, E. Chason, L.B. Freund, M. Sinclair, R.D. Twesten, and R.Q. Wang, *Phys. Rev. Lett.* **80**, 4717 (1998).

[49] F. Montalenti, P. Raiteri, D.B. Migas, H. von Känel, A. Rastelli, C. Manzano, G. Costantini, U. Denker, O.G. Schmidt, K. Kern, and L. Miglio, *Phys. Rev. Lett.* **93**, 216102 (2004).

[50] M. Brehm, F. Montalenti, M. Grydlik, G. Vastola, H. Lichtenberger, N. Hrauda, M.J. Beck, T. Fromherz, F. Schäffler, L. Miglio, and G. Bauer, *Phys. Rev. B* **80**, 205321 (2009).

[51] X. Chen, F. Wu, Z. Zhang, and M.G. Lagally, *Phys. Rev. Lett.* **73**, 850 (1994).

[52] F. Wu and M.G. Lagally, *Phys. Rev. Lett.* **75**, 2534 (1995).

[53] A. Rastelli, H. von Känel, G. Albini, P. Raiteri, D.B. Migas, and L. Miglio, *Phys. Rev. Lett.* **90**, 216104 (2003).

[54] D.B. Migas, P. Raiteri, L. Miglio, A. Rastelli, and H. von Känel, *Phys. Rev. B* **69**, 235318 (2004).

[55] M.J. Beck, A. van de Walle, and M. Asta, *Phys. Rev. B* **70**, 205337 (2004).

[56] J. Tersoff, *Phys. Rev. B* **43**, 9377(R) (1991).

[57] F. Zipoli. S. Cereda, M. Ceriotti, M. Bernasconi, L. Miglio, and F. Montalenti, *Appl. Phys. Lett.* **92**, 191908 (2008).

[58] E. Bussmann and B.S. Swartzentruber, *Phys. Rev. Lett.* **104**, 126101 (2010).

[59] M. Brehm, M. Grydlik, H. Lichtenberger, T. Fromherz, N. Hrauda, W. Jantsch, F. Schäffler, and G. Bauer, *Appl. Phys. Lett.* **93**, 121901 (2008).

[60] I. Daruka and A.-L. Barabási, *Phys. Rev. Lett.* **79**, 3708 (1997).

[61] G.-H. Lu and F. Liu, *Phys. Rev. Lett.* **94**, 176103 (2005).

[62] T. Miyazaki, D.R. Bowler, M.J. Gillan, and T. Ohno, *J. Phys. Soc. Jpn* **77**, 123706 (2008).

[63] A.A. Stekolnikov and F. Bechsted, *Phys. Rev. B* **72**, 125326 (2005).

[64] R. Marchetti, F. Montalenti, L. Miglio, G. Capellini, M. De Seta, and F. Evangelisti, *Appl. Phys. Lett.* **87**, 261919 (2005).

[65] D. Digiuni, R. Gatti, and F. Montalenti, *Phys. Rev. B* **80**, 155436 (2009).

[66] F. Arciprete, E. Placidi, V. Sessi, M. Fanfoni, F. Patella, and A. Balzarotti, *Appl. Phys. Lett.* **89**, 041904 (2006).

[67] L.G. Wang, P. Kratzer, M. Scheffler, and N. Moll, *Phys. Rev. Lett.* **82**, 4042 (1999).

[68] P. Kratzer, Q.K.K. Liu, P. Acosta-Diaz, C. Manzano, G. Costantini, R. Songmuang, A. Rastelli, O.G. Schmidt, and K. Kern, *Phys. Rev. B* **73**, 205347 (2006).

[69] A. Madhukar, *Surf. Sci.* **132**, 344 (1983).

[70] M. Meixner, R. Kunert, and E. Schöll, *Phys. Rev. B* **67**, 195301 (2003).

[71] G. Russo and P. Smereka, *J. Comput. Phys.* **214**, 809 (2006).

[72] C.-H. Lam, M.T. Lung, and L.M. Sander, *J. Sci. Comput.* **37**, 73 (2008).

[73] B.J. Spencer, P.W. Voorhees, and J. Tersoff, *Phys. Rev. B* **64**, 235318 (2001).

[74] Y.W. Zhang, *Phys. Rev. B* **61**, 10388 (2000).

[75] P. Liu and Y.W. Zhang, *Int. J. Solids Struct.* **44**, 1733 (2007).

[76] A. Malachias, S. Kycia, G. Medeiros-Ribeiro, R. Magalhães-Paniago, T.I. Kamins, and R.S. Williams, *Phys. Rev. Lett.* **91**, 176101 (2003).

[77] G. Medeiros-Ribeiro and R.S. Williams, *Nano Lett.* **7**, 223 (2007).

[78] T.U. Schülli, G. Vastola, M.-I. Richard, A. Malachias, G. Renaud, F. Uhlík, F. Montalenti, G. Chen, L. Miglio, F. Schäffler, and G. Bauer, *Phys. Rev. Lett.* **102**, 025502 (2009).

[79] U. Denker, M. Stoffel, and O.G. Schmidt, *Phys. Rev. Lett.* **90**, 196102 (2003).

[80] G. Katsaros, G. Costantini, M. Stoffel, R. Esteban, A.M. Bittner, A. Rastelli, U. Denker, O.G. Schmidt, and K. Kern, *Phys. Rev. B* **72**, 195320 (2005).

[81] A. Rastelli, M. Stoffel, A. Malachias, T. Merdzhanova, G. Katsaros, K. Kern, T.H. Metzger, and O.G. Schmidt, *Nano Lett.* **8**, 1404 (2008).

[82] A. Baskaran, J. Devita, and P. Smereka, *Continuum Mech. Thermodyn.* (2009).

[83] J. Tersoff, *Appl. Phys. Lett.* **83**, 353 (2003).

[84] F. Uhlík, R. Gatti, and F. Montalenti, *J. Phys.: Condens. Matter* **21**, 084217 (2009).

[85] B.J. Spencer and M. Blanariu, *Phys. Rev. Lett.* **95**, 206101 (2005).

[86] C. Lang and D.J.H. Cockayne, *Phys. Rev. B* **72**, 155328 (2005).

[87] G. Hadjisavvas and P.C. Kelires, *Phys. Rev. B* **72**, 075334 (2005).

[88] R. Gatti, F. Uhlík, and F. Montalenti, *New J. Phys.* **10**, 083039 (2008).

[89] N.V. Medhekar, V. Hegadekatte, and V.B, Shenoy, *Phys. Rev. Lett.* **100**, 106104 (2008).

[90] R. Phillips, *Crystals, Defects and Microstructures*, Cambridge University Press, Cambridge, (2001).

[91] J.D. Eshelby, *Proc. R. Soc. Lond. A* **241**, 376 (1957).

[92] L. Huang, G.-H. Lu, F. Liu, and X.G. Gong, *Surf. Sci.* **601**, 3067 (2007).

[93] J. Tersoff, C. Teichert, and M.G. Lagally, *Phys. Rev. Lett.* **76**, 1675 (1996).

[94] J. Tersoff, *Phys. Rev. B* **39**, R5566 (1989).

[95] F. Montalenti, A. Marzegalli, G. Capellini, M. De Seta, and L. Miglio, *J. Phys.: Condens. Matter* **19**, 225001 (2007).

[96] A.R. Woll, P. Rugheimer, and M.G. Lagally, *Mat. Sci. Eng. B* **96**, 94 (2002).

[97] Q. Xie, A. Madhukar, P. Chen, and N.P. Kobayashi, *Phys. Rev. Lett.* **75**, 2542 (1995).

[98] A.A. Maradudin and R.F. Wallis, *Surf. Sci.* **91**, 423 (1980).

[99] H.X. Zhong, J.C. Wells, Q. Liu, and Z.Y. Zhang, *Surf. Sci.* **539**, L525 (2003).

[100] M.A. Makeev and A. Madhukar, *Phys. Rev. Lett.* **86**, 5542 (2001).

[101] A. Rastelli, M. Kummer, and H. von Känel, *Phys. Rev. Lett.* **87**, 256101 (2001).

[102] M. De Seta, G. Capellini, and F. Evangelisti, *Phys. Rev. B* **71**, 115308 (2005).

[103] X. Tan, X.L. Li, and G.W. Yang, *Phys. Rev. B* **77**, 245322 (2008).

[104] G. Capellini, M. De Seta, F. Evangelisti, V.A. Zinovyev, G. Vastola, F. Montalenti, and L. Miglio, *Phys. Rev. Lett.* **96**, 106102 (2006).

[105] U. Denker, A. Rastelli, M. Stoffel, J. Tersoff, G. Katsaros, G. Costantini, K. Kern, N.Y. Jin-Phillipp, D.E. Jesson, and O.G. Schmidt, *Phys. Rev. Lett.* **94**, 216103 (2005).

[106] Z. Zhong and G. Bauer, *Appl. Phys. Lett.* **84**, 1922 (2004).

[107] F. Pezzoli, M. Stoffel, T. Merdzhanova, A. Rastelli, and O.G. Schmidt, *Nanoscale Res. Lett.* **4**, 1073 (2009).

[108] J. Zhang, F. Montalenti, A. Rastelli, N. Hrauda, D. Scopece, H. Gross, J. Stangl, F. Pezzoli, F. Schäffler, O.G. Schmidt, and G. Bauer *Phys. Rev. Lett.*, **105**, 166102, (2010).

[109] F. Boioli *et al.*, *in preparation.*

11

Strain engineering of silicon–germanium (SiGe) micro- and nanostructures

F. PEZZOLI, C. DENEKE and O. G. SCHMIDT,
IFW Dresden, Germany

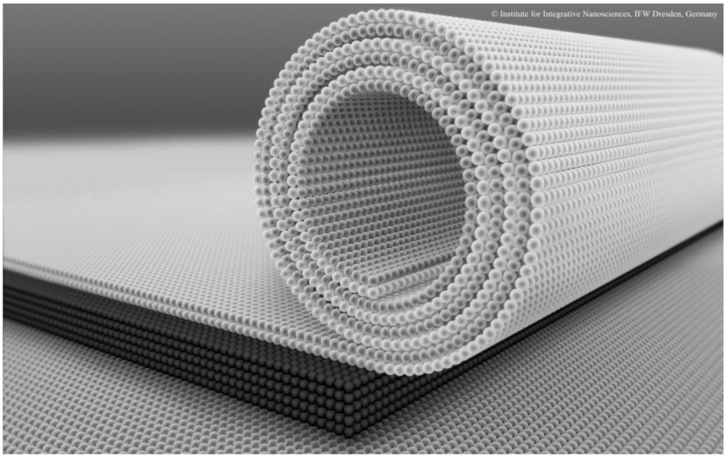
Abstract: This chapter is intended to provide an overview of strain-engineered heterostructures and quantum devices based on SiGe alloys. The growth of SiGe on Si(001) substrates is introduced by focusing on the morphological evolution of SiGe nanostructures and the ways of precisely controlling lateral and vertical ordering. Afterwards, the chapter discusses a revolutionary process technology leading to strain-driven architectures. Finally, a new emerging generation of SiGe-based systems with unique capabilities, ranging from fast field-effect transistors to energy harvesting devices, is reviewed.

Key words: SiGe alloys, strain engineering, semiconductor nanostructures, self-organization, Ge islands.

11.1 Introduction

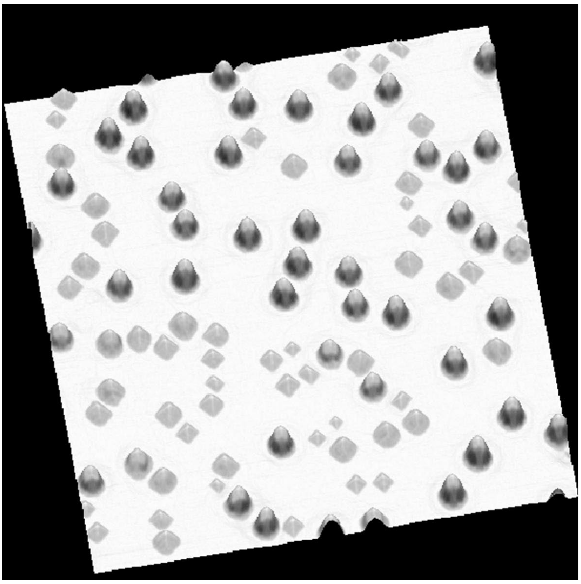
The study of phenomena based on stress generation on surfaces and in materials is currently one of the hottest topics in semiconductor physics, because it provides insight and guidelines for the control of sophisticated structures at the nanometer scale. Archetypal strain-driven architectures (Figs 11.1 and 11.2) are highly strained layers which roll up into tubular structures, once the films are released from the substrate surface (Prinz *et al*., 2000; Schmidt and Eberl, 2001a), or three-dimensional (3D) islands, namely quantum dots (QD), spontaneously forming during lattice mismatched epitaxy (Shchukin and Bimberg, 1999). So-called *strain engineering*, which consists of fine tuning of the strain effects in nanostructures, is particularly important not only as a means to fabricate low-dimensional structures, but also opening up new degrees of freedom via the almost arbitrary modification of the semiconductor energy bands (People and Bean, 1986). This in turn allows a deterministic and precise definition of the electronic and optical properties of a material, enabling the exploitation of new and innovative functionalities (Capasso, 1987).

A prominent approach to designing strain-induced architectures is based on the epitaxial growth of structures comprising alternating layers of two or more materials with different equilibrium lattice constants: heteroepitaxy



© Institute for Integrative Nanosciences, IFW Dresden, Germany

11.1 Schematic view of a micro- or nanotube incorporating different materials.



11.2 2.5 × 2.5 μm^2 AFM image of Ge nanostructures on Si(001). The grayscale is related to the surface slope ranging from 0 to 40° with respect to the Si(001) growth plane.

(Narayanamurti, 1987). The establishment and constant development of accurate, reliable and well-controlled growth techniques has been therefore one of the key factors of the extremely exciting advances in nanotechnology.

It is illuminating to note that, for instance, a transistor based on SiGe alloys was proposed already in the 1950s (see Paul, 2004, and reference therein), but, due to the difficulty in growing high-quality SiGe substrates, in practice such a device was not demonstrated until 1975, when epitaxial growth of SiGe/Si was employed (Kasper *et al.*, 1975). Nowadays, growth techniques have finally allowed the controlled fabrication of structures at a nanoscale level, initiating intriguing investigations into low-dimensional physics (Shchukin and Bimberg, 1999; Teichert, 2002). For instance, the past several years have seen remarkable progress towards the realization of quantum photonics. Carrier confinement in nanostructures is indeed expected to favor room-temperature applications and to increase the efficiency of radiative recombination processes. The growth of QDs has already enabled quantum light generation based on single-photon sources (Shields, 2007; Michler *et al.*, 2000). As another example, nanostructuring of semiconductors is one of the most promising and elegant approaches to increase thermoelectric performance. In recent reports on Si nanowires (Boukai *et al.*, 2008; Hochbaum *et al.*, 2008), the reduction of dimensions at the nanometer scale demonstrated how a semiconductor such as Si turns out to be appealing for thermoelectrics, albeit traditionally considered an uninteresting material for such applications.

So far, Ge on Si(001) has emerged as a model system to study growth physics in lattice mismatched heterostructures, because it is formed by fully miscible non-polar semiconductors and its phenomena can be easily mapped onto other material systems (Costantini *et al.*, 2004). Besides, it combines the nanotechnology benefits with monolithic integration into the mainstream Si technology (Paul, 2004; Schäffler, 1997). During the 1980s, several discoveries about the band alignments in SiGe heterostructures allowed fundamental properties of crystal structures known as *superlattices* and *quantum wells* to be exploited (Schäffler, 1997). More recently, conceptually new top-down approaches to creating Si-based heterostructures and superlattices have been suggested (see, for instance, Li, 2008, and Deneke *et al.*, 2009). By the release and roll-up of layer systems, nano- and microtubes are formed, which consist of walls that incorporate alternating layers of SiGe/Si, SiO_2/Si or even otherwise incompatible material classes such as semiconductor/organic or semiconductor/metal superlattices, e.g. Deneke *et al.* (2009). The walls of these radial superlattices can be designed in such a way that they become relevant both fundamentally as well as application-wise (Deneke *et al.*, 2008). The method has gained considerable attention, due to the high potential these tubes possess in interdisciplinary fields (Schmidt and Eberl, 2001a; Golod *et al.*, 2001). In particular, these tubes have been suggested as nanopipelines for fluid transportation and fluid storage on substrate surfaces (Schmidt and Eberl, 2001a; Schmidt and Jin-Phillipp, 2001). In addition, by compressing the tubes it is even possible to transform the radial superlattice structures into planar superlattices, demonstrating large design flexibility (Zander *et al*., 2009).

On the other hand, the 4.2% lattice mismatch between pure Si and pure Ge is one of the main driving forces leading to the self-assembly of islands during strained layer growth (Brunner, 2002). At an early stage of nanostructure research, SiGe heteroepitaxy in the regime favoring non-planar growth has been recognized as an attractive bottom-up approach to synthesizing novel structures at the nanometer scale (Eaglesham and Cerullo, 1990; Mo *et al*., 1990). Considerable work has been done on growth and characterization of SiGe islands, and important results have been summarized in several reviews (Brunner, 2002; Stangl *et al*., 2004; Baribeau *et al*., 2006; Berbezier and Ronda, 2009). In this chapter we briefly outline progress in the growth, discussing the evolution of the island morphology with the growth parameters, and stressing the importance of the Ge coverage and the substrate temperature (Section 11.2). We examine recent experimental results pertaining to the control of the density and size, by describing the motivations and approaches to obtaining (i) vertical alignment; (ii) lateral ordering; and (iii) how these ordering processes can be modified and controlled (Section 11.3). Complementarily, in Section 11.4, we discuss the 3D fabrication based on a stress-driven actuation principle, emphasizing integrative rolled-up SiGe-based tubes. We look at some issues and ideas in stress-driven architectures as a new path to fabricating CMOS-compatible Si nano-devices, such as field-effect transistors from strained Si bridges (Section 11.5). We also address exploratory research into SiGe islands within the framework, for instance, of thermoelectrics. In addition, this section provides an outlook on latest developments concerning stress-driven architectures such as rolled-up electronics. Finally Section 11.6 provides a short commentary on research groups, websites, etc., related to the themes of the chapter.

11.2 Growth insights

11.2.1 Heteroepitaxy

Epitaxy can be defined as a well-controlled phase transition which leads to a single crystalline solid (Stringfellow, 1982). In this process a crystalline phase, the *epilayer*, is forced to grow in a structure-dependent manner onto another crystalline phase, the *substrate*, of a given structure. At the substrate surface, different atomic processes take place, leading to a chemically and structurally inhomogeneous interfacial region. Epilayers grown in a single-component system, that is, with the same chemical composition as the substrate, are called *homoepitaxial* layers, even if they differ in doping with electrically active impurities. Homoepitaxial layers can be deposited under conditions in which the crystallographic orientation of the layer is exactly determined by the substrate. In contrast, *heteroepitaxy* is the growth of a crystalline film on a crystalline substrate consisting of a different material. In this case, the deposition process is affected by the mutual relations between the substrate and the heteroepilayer (Herman, 1999). Differences in atomic spacing of the two lattices, in the absence of interfacial interaction between them, are responsible for the misalignment of the equilibrium interfacial atomic arrangements. As a consequence, the deposition process depends strongly on whether the grown epilayer is coherent or incoherent with the substrate, i.e. on whether the interface between the epitaxial overgrowth and the substrate is crystallographically perfect or not (Wood and Zunger, 1989). When the epilayer is incoherent with the substrate, it is free to adjust the in-plane lattice constant to minimize its free energy. If, on the other hand, it is coherent, the epilayer energy is minimized by adopting the in-plane lattice constant of the substrate, thus creating a commensurate material system with the substrate. The resulting strain in the epilayer can then increase its overall free energy significantly. Further crucial problems that influence the particularities of the growth are the crystallographic orientation of the substrate and its surface reconstruction. Finally, a minor role in the growth process is played by the thermal stress, the defects appearing at the crystal–film interface, and the chemical interaction between the epilayer and the substrate, including segregation of the substrate elements towards the free surface.

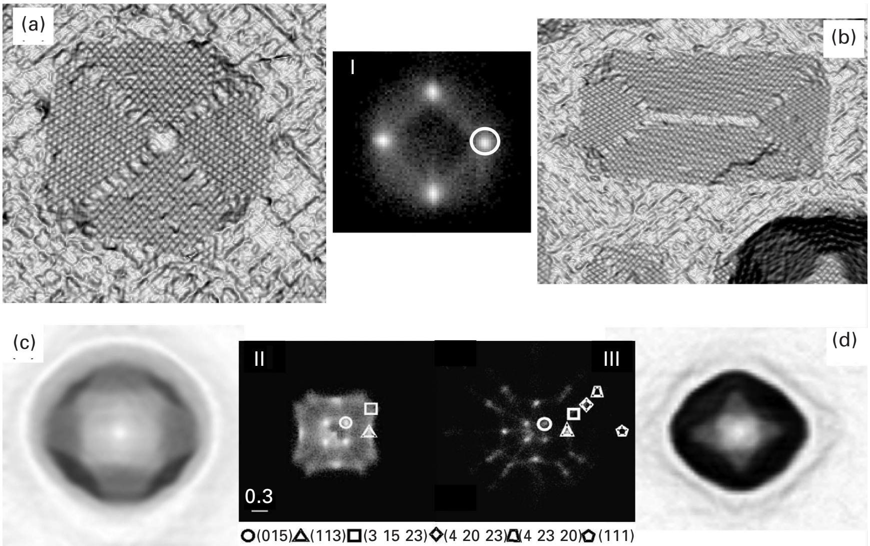
In the equilibrium theory of heteroepitaxy, three growth modes are distinguished (Shchukin and Bimberg, 1999): (i) Frank–van der Merwe (FM); (ii) Volmer–Weber (VW); and (iii) Stranski–Krastanow (SK). These can be described as layer-by-layer growth (2D), island growth (3D) and layer-by-layer (wetting layer, WL) plus island growth, respectively. The particular growth mode for a given system depends on the interface energies and on the lattice mismatch. Since the SiGe surface energy is lower than that of Si, two-dimensional growth would be favored in SiGe/Si heteroepitaxy. However, due to the lattice mismatch and the small interface energy, for alloy compositions with a Ge content greater than about 10%, plastic relaxation is usually preceded by the formation of coherently strained 3D morphologies (Brunner, 2002). Thus at high misfit regimes, the initial growth may occur layer-by-layer, but as the layer thickness increases, the strain energy increases as well. This continues until the system lowers its energy by forming isolated islands in which strain is relaxed.

A detailed description of SK growth mode has been given in Chapter 10. In the early days of crystal growth, bidimensional, flat strained epilayers were the basic constraint for the fabrication of devices based on strain engineering. Growth conditions leading to island nucleation, although already reported in the late 1960s and early 1970s (Ito and Takahashi, 1968; Cullis and Booker, 1971), were generally avoided, due to the detrimental effects on device performance. Only recently has heteroepitaxy, in the regime favoring island growth, been recognized as an interesting bottom-up approach to fabricating nanostructures (Nötzel *et al*., 1994). Indeed, in 1990 Eaglesham and Cerullo showed by means of transmission electron microscopy (TEM) that Ge islands grown on Si(001) at 500°C are free from defects, namely dislocations, up to a thickness of about 50 nm. Remarkably, this height exceeds by a factor of 50 the critical thickness above which dislocations start to appear in flat films (Eaglesham and Cerullo, 1990). This can only be explained by the elastic relaxation of the island and of the surrounding area of the substrate. Those findings have generated considerable interest and stimulated detailed investigations concerning the mechanisms of island formation and shape tailoring, paving the way to the fabrication of ordered and defect-free arrays of self-assembled quantum-confined structures based on Si, which are known to have novel electronic and optoelectronic applications (Wang *et al*., 2005; Tsybeskov and Lockwood, 2009). Finally, by optimizing growth parameters, it has been shown that it is possible to fabricate such self-organized semiconductor nanostructures with properties optimized for the desired application (Berbezier *et al*., 2002; Capellini *et al*., 2003).

It should be noted that the SiGe nanostructure formation process is driven by thermodynamics, i.e. strain relief overcompensates for the increase in surface energy associated with the island development. As a consequence, the resulting nanostructures might be metastable and their exact shape, size and composition result from a delicate interplay between thermodynamic as well as kinetic effects. In addition, the geometry of these self-organized nanostructures is also dictated by the crystalline orientation of the Si surface (Persichetti *et al*., 2010), leading to a large variety of shapes and morphologies. Such a subject is beyond the scope of this review, and we will therefore restrict ourselves to the phenomena related to the growth on the Si(001) surface. Comprehensive reviews concerning the growth mechanisms on vicinal surfaces and on Si(111) have been recently provided by Motta *et al*. (2003) and Berbezier and Ronda (2009).

11.2.2 Island evolution

In the following we focus on the morphological evolution of islands and their transitions to steeper morphologies as their size increases (Rastelli *et al*., 2006; Berbezier *et al*., 2003). The Ge overlayers on Si(001) grow layer by layer up to 3–5 monolayers (1 ML = 6.78×10^{14} Ge atoms/cm^2). The coherent Ge wetting layer of a thickness of about 2 ML already shows a quasiperiodic strain-induced $(2 \times n)$ missing dimer structure, which becomes an $(m \times n)$ structure at 3 ML (Köhler *et al*., 1992; Voigtländer, 2001; Tomitori *et al*., 1994). At higher Ge coverage two different scenarios are commonly observed in the experiments. During low lattice misfit heteroepitaxy, shallow unfaceted mounds, called prepyramids, have been reported (Vailionis *et al*., 2000; Sutter and Lagally, 2000; Tromp *et al.*, 2000; Tersoff *et al.*, 2002). This strain-induced roughening of an epitaxial film is known as the *Asaro–Tiller–Grinfeld* (ATG) instability (Asaro and Tiller, 1972; Grinfeld, 1994 and references therein; Srolovitz, 1989). From a thermodynamic point of view, a perturbation at the surface of a planar strained layer allows lattice planes to relax towards the ripple tips. This strain-relief mechanism by roughening (Grinfeld, 1994; Srolovitz, 1989; Horn-von Hoegen *et al.*, 1994) has been quantified in film stress measurements by Schell-Sorokin and Tromp (1990) and Wedler *et al.* (1998). This reduction of the elastic energy is, however, accompanied by an increase of the surface energy. These phenomena occur because of the adatom surface diffusion towards sites with the lowest free energy. For SiGe on Si, ATG instability evolves from prepyramids to islands with well-developed crystallographic facets, namely pyramids (shown in Fig. 11.3(a)). Theoretical calculations (Tersoff *et al.*, 2002), supported by STM analysis (Rastelli *et al.*, 2003), showed that faceted pyramids are the resulting stable shape of a first-order shape transition of the mounds, where at first small {1 0 5} facets (see Fig. 11.3(I)) nucleate at regions of the prepyramid

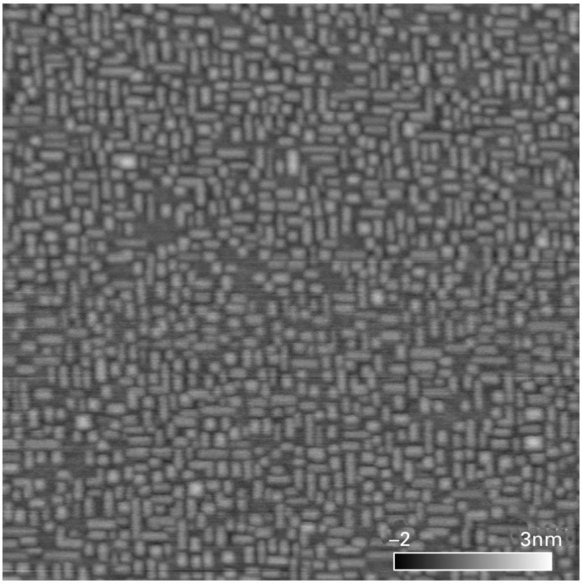


11.3 (a) and (b): STM images (about 30 nm × 30 nm) of a pyramid and a hut cluster, respectively. (c) and (d): AFM images of a dome (about 500 nm × 500 nm) and a barn (about 300 nm × 300 nm), respectively. The contrast is proportional to the local surface slope with respect to the (001) Si surface plane, thus allowing the identification of shallow facets (light regions from 0°) and steep facets (dark regions up to 45°). (I)–(III): Histograms of the surface gradient, the so-called facet plots, averaged on a large number of islands according to Rastelli and von Känel (2002).

surface located at a height intermediate between the base and the top, then, due to their low surface tension (Raiteri *et al.*, 2002; Fujikawa *et al.*, 2002), the {1 0 5} facets expand in order to reduce the surface energy, and the unfaceted remnants of the parent prepyramid disappear. The pyramid apex forms, because being partially relaxed it collects deposited adatoms from the surrounding area. Finally the pyramid ripens rather self-similarly and the shallow mound at the island base is replaced by a trench.

On the other hand, for large misfit values the Stranski–Krastanow growth mode plays a major role. The development of a 3D morphology is a self-organization mechanism driven by the minimization of the system energy. This is obtained by means of island nucleation, which allows the relief of elastic energy. The process, however, is always accompanied by an increase of the surface energy, which depends strongly on the formation of various crystallographic facets. The energy change dependence on island volume shows an increase to a maximum value, E_C, which represents a thermal activation energy for island nucleation. E_C depends on the facet orientation and it has been demonstrated that it scales with the strain, ε, as ε^{-4}, or with the Ge content as x^{-4}, rendering the transition to the SK growth mode in low-strain systems unfavorable, in contrast with the ATG instabilities (see, for instance, Brunner, 2002; Baribeau *et al.*, 2006 and references therein). In this regime, at a coverage exceeding the wetting layer thickness of 3–5 ML, coherent nanocrystals in the shape of prismlike 'huts' are commonly observed in a low-temperature range, e.g. between 300 and 600°C. An AFM image of a hut cluster ensemble obtained upon deposition of 5.8 ML of Ge at 500°C is shown in Fig. 11.4. As highlighted in Figs 11.3(b) and 11.3(I), these nanostructures exhibit {1 0 5} facets, such that the contact angle with the Si(001) substrate is about 11° (Tomitori *et al.*, 1994). These regularly shaped hut clusters were first described in 1990 by Mo and Lagally (Mo *et al.*, 1990), who explained them as one step on the kinetic pathway to the formation of large 3D clusters. The thickness fluctuations of the Ge wetting layer act as a means for the formation of the hut clusters (Goldfarb *et al.*, 1997; Jesson *et al.*, 2000), allowing the strain relief of Ge through an elastic distortion of the coherent lattice. This decrease in energy outbalances the higher surface free energy of the {1 0 5} facets leading to the facet formation, without dislocation injection.

The kinetics of the nucleation and the formation of the huts have been studied by means of several techniques such as scanning tunneling microscopy (STM) and low energy electron diffraction (LEED) (Goldfarb *et al.*, 1997; Kästner and Voigtländer, 1999; Horn-von Hoegen *et al.*, 2004). Medeiros-Ribeiro *et al.* (1998a) demonstrated that hut clusters are metastable. Hut clusters are indeed usually observed only in samples grown at low temperature and tend to disappear during annealing experiments, leaving the surface populated either by small {1 0 5}-faceted pyramids with square

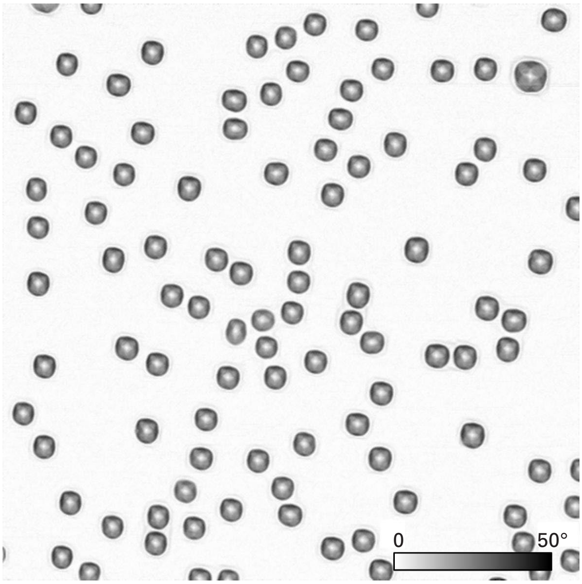


11.4 AFM image (1 μm × 1 μm) of hut clusters obtained by MBE deposition of 5.8 ML Ge on Si(001) at 500°C. Here the grayscale is related to the surface slope with respect to the (001) plane.

base or by larger, multifaceted islands. It can be concluded that the shape of the hut clusters must have a kinetic origin and the equilibrium shape of {1 0 5} faceted islands is that of pyramids. Besides, Voigtländer (2001) studied *in vivo* the evolution of hut clusters by STM, clearly showing that the elongation is due to an anisotropy in the growth rate of facets favoring their lateral expansion. By assuming that the islands are composed of small (001) terraces and steps, the barrier for the nucleation of a new layer on a {1 0 5} facet increases with increasing facet size, thus favoring the growth of the smallest facet and leading to the observed elongation, as previously suggested by Jesson *et al.* (1998). Nevertheless, more recently McKay *et al.* (2008) showed that an ensemble of huts can be kinetically stabilized during annealing once the areal density is sufficiently low, i.e. ripening is suppressed since there are no lower chemical potential islands.

As the size of the dot increases, a transition towards facets steeper than the {1 0 5} occurs, finally leading to islands with larger ratios of diameter and height. This is an important structural parameter usually called the *aspect ratio*, abbreviated to ar. The metastable hut clusters (ar < 0.2) decompose towards so-called dome islands, characterized by 0.2 < ar < 0.28 (Medeiros-Ribeiro *et al.*, 1998b). This sequence of shapes has been observed for the growth of SiGe on Si(001) at least for Ge molar fractions larger than 0.2 (Floro *et al.*, 1998a,b). The transition from hut to dome clusters has been observed in a real-time LEEM investigation by Ross *et al.* (1999), and has been explained as a first-order transition by Daruka *et al.* (1999) and Costantini *et al.* (2004). As reported in Fig. 11.3(c) and in the corresponding facet plot Fig. 11.3(II), domes are octagonal-based islands with a body bounded by four {113} with an inclination of 25°, and eight {15 3 23}, with an inclination of 33°. {1 0 5} facets are also present at the island apex and foot. They appear to be a stable part of the dome shape, since they are also observable even after annealing in Si-rich domes. The role of shallow {1 0 5} facets at the dome base has been recognized by Daruka and Tersoff (2002) and introduced in their theoretical model of island shape evolution (Daruka *et al.*, 1999). Though a dome is larger than a hut and contains much more Ge, it is still coherent (Medeiros-Ribeiro *et al.*, 1998b; Eaglesham and Cerullo, 1990): the strain relief via dome clusters is much more efficient due to the steeper facets. In a narrow substrate temperature and Ge coverage window a highly uniform distribution of dome-shaped islands has been reported by several authors. See, for instance, Fig. 11.5, where a monomodal ensemble of domes has been obtained by growing 6.7 ML Ge at 620°C. In particular, extremely small height deviations of about ±3% and ±5.5% were found by Jin *et al.* (2003) and Schülli *et al.* (2005), respectively.

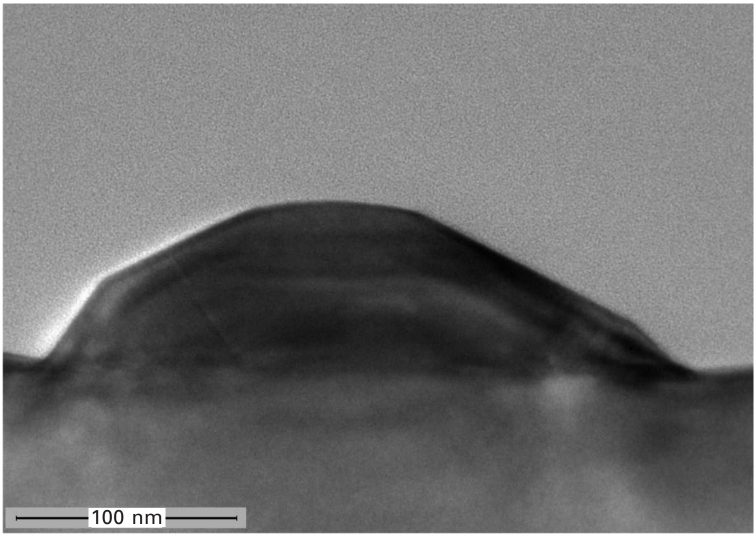
Surprisingly, during the deposition of SiGe on Si(001) with $x < 0.2$, Sutter



11.5 AFM image (2 μm × 2 μm) of SiGe islands obtained after deposition of 6.7 ML Ge on Si(001) at 620°C. The shading allows the various facets to be distinguished according to their local surface slope with respect to the (001) plane.

et al. (2004) have identified islands bounded by steep facets, namely {1 1 1}, which are not commonly observed in domes. Such a new type of coherent island, the so-called 'barn' (having ar ≈ 0.3 and shown in Fig. 11.3(d)), has been reported also during heteroepitaxial growth of pure Ge on Si(001) in the temperature range from 620°C to 750°C (Stoffel *et al.*, (2006). By means of a detailed morphological AFM investigation, Stoffel *et al.* (2006) demonstrated that, similarly to the pyramid-to-dome transition, domes evolve towards barns via the appearance of much steeper facets which grow at the expense of shallower ones. In this case the island base is bounded by {1 1 1} and in addition by {20 4 23} facets (see Fig. 11.3(III)). The presence of the latter, corresponding to a major stable surface for Si but not for Ge, suggests that a strong intermixing takes place during the dome–barn transition. This led to an extension of the previously known sequence of equilibrium shapes for strained Ge nanocrystals on Si and proved the existence of a delay of the plastic relaxation onset. Interestingly, the barrier for dislocation introduction stabilizes a monodisperse island size distribution: a barn height distribution having a very narrow full-width at half-maximum of about ±6.7% was reported for a Ge coverage of 8.5 ML deposited at 700°C (Stoffel *et al.*, 2006).

Nevertheless, for high Ge coverages, plastic relaxation spontaneously occurs within the island itself, as shown in Fig. 11.6 (Marzegalli *et al.*, 2007). LeGoues and co-workers were among the first to demonstrate by means of



11.6 Cross-sectional transmission electron micrograph in the [110] projection of a superdome. Dislocation threading arms propagating through the nanostructure can be better observed on the left side of the island.

real-time and in-situ TEM that the growth mode of large dislocated islands (superdomes) is 'cyclic' in the temperature range between 350 and 650°C, i.e. islands quickly grow after the introduction of a new dislocation (LeGoues *et al.*, 1994, 1996; Hammar *et al.*, 1996; Liu *et al.*, 2000). But, despite superdomes being commonly observed in a wide temperature range, so far only a few papers have addressed this topic (see, for instance, LeGoues *et al.*, 1994; Hammar *et al.*, 1996; Liu *et al.*, 2000). By using grazing-incidence small-angle X-ray scattering (GISAXS) and diffraction (GIXD), Richard *et al.* (2009) investigated *in situ* the evolution of superdomes deposited at very low growth rates of 0.006 ML/s, and relatively high temperatures between 770 and 920°C. Within these growth conditions, the onset for dislocation formation is observed at lower Ge coverages than in previous reports. In addition, coalescence and anomalous coarsening have been shown to play a major role in the transition towards plastic relaxation, which has been suggested as a mechanism leading to the stabilization of the islands in terms of their average lattice parameter.

Recently, a combination of selective wet chemical etching and scanning probe microscopy has been shown to be a much easier technique, compared to X-ray and TEM, to investigate the evolution of the relaxed islands (Merdzhanova *et al.*, 2006). The analysis of the pattern, which consists of ring structures left on the Si surface after island removal, has been used to gather information on the evolution of the superdomes and their interaction with the surrounding material (Merdzhanova *et al.*, 2006; Denker *et al.*, 2001). This work clarified the results of LeGoues *et al.* (1994), showing that the observed rings correspond to the formation of a new trench when a dislocation is introduced in the island.

From the theoretical point of view, only recently Gatti *et al.* (2008) developed an *ad hoc* methodology for numerical calculation in the framework of the finite element method (FEM). This approach enables the exact calculation of the elastic field due to dislocations embedded in a finite structure, by considering the dislocation field as it is in a bulk system plus a correction term due to the presence of free surfaces. The procedure has already been applied successfully to a straight dislocation segment in a Ge island. However, little work has been done so far to understand the mechanism of nucleation and interaction of the dislocations within the SiGe superdomes and some phenomena have not yet been understood, i.e. (i) how the first ring is introduced; (ii) the actual shape of the dislocation loops (round, polygonal, etc.), (iii) why a dislocation is created periodically as the island grows, and (iv) the role of the repulsion between parallel dislocations.

A crucial problem related to the self-assembled island growth is the influence on morphology, strain and chemical composition of the growth parameters and post-growth treatments. It has been shown that by increasing temperature, the average size of the Ge islands grown by MBE increases

(Chaparro *et al.*, 2000) and the size distribution becomes narrower (Drucker and Chaparro, 1997). Then, in Ge islands, the Si overgrowth triggers a sequence of shape transition nearly the reverse of those during the growth (Rastelli *et al.*, 2001). This shape reversal was explained by the incorporation of Si into the island (Costantini *et al.*, 2004). As a consequence, the overgrowth of Ge domes with Si causes changes in the composition and the strain state. It has been shown that the intermixing during overgrowth can be suppressed, preserving island features, by using a low growth temperature, typically below 400°C (Rastelli *et al.*, 2001, 2002; Schmidt *et al.*, 2000a; Hesse *et al.*, 2002; Capellini *et al.*, 2005; De Seta *et al.*, 2002), even though low-temperature capping might eventually lead to the formation of planar defects in the Si layer, as recently pointed out by Lin *et al.* (2009). The intermixing between the islands and their immediate neighbors during growth and overgrowth has been investigated in detail using time-resolved transmission electron microscopy and high-resolution X-ray diffraction (Hesse *et al.*, 2002; Lang *et al.*, 2006). All the results outlined above have been studied extensively and the key findings are summarized in the review articles by Stangl *et al.* (2004), Baribeau *et al.* (2006) and Rastelli *et al.* (2006).

In conclusion, in SiGe/Si heteroepitaxy the strain is one of the main driving forces determining the final layer morphology. SK growth mode, ATG instability and dislocation injection are indeed competing strain relief mechanisms: the critical nucleus size for Stranski–Krastanow growth scales with the lattice mismatch, ε, as ε^{-4} (Tersoff and LeGoues, 1994), whereas the ATG morphological instability scales as ε^{-2} (see Stangl *et al.*, 2004, and references therein), and the plastic relaxation via misfit dislocations depends on ε^{-1} (Tersoff and LeGoues, 1994). As a consequence, ATG roughening and misfit dislocation nucleation are effective in low-strain regimes; in particular, the plastic relaxation is favored only below a given stress value (Tersoff and LeGoues, 1994; De Seta *et al.*, 2006). Nevertheless, during growth, kinetic factors (Mo *et al.*, 1990) and surface effects, such as the growth rate, the substrate temperature and the presence of surfactants, might suppress surface diffusion, reducing the roughness and facilitating the planarization of the surface of the strained layer, even if the epilayer is in a high-stress regime. Finally, it should be noted that despite nearly 20 years of extensive investigations on Ge islands, the observations of self-assembly processes still cannot completely cover the richness of phenomena taking place on the well-known Si(001) surface. Island morphologies and final ensemble configurations are specific outcomes of a delicate counterbalance between various and even complex mechanisms. Very recently, Brehm *et al.* (2009) demonstrated that also the WL might define the shape of stable islands. An accurate and high resolution (0.025 ML) control of the WL thickness led to unexpected observations about the early stage of the island formation. For samples grown at 700°C and WL thickness in the range from 4.32 to 4.38

ML, it was found that domes can nucleate before pyramids. This trend seems to take place for temperatures above 675°C. Brehm *et al.* (2009) introduced a thermodynamic model to clarify those findings, demonstrating that below a certain WL thickness, domes are stable islands whereas pyramids are not, which goes along with a WL thinning followed by dome nucleation at temperature-dependent overcritical thicknesses.

11.3 Island engineering

In the previous section we outlined experimental evidence where SiGe islands have well-defined shapes and morphologies. But, since the temperature plays an important role in defining adatom diffusion statistics during heteroepitaxial growth, those findings are always characterized by a spontaneous, random nucleation of the islands. In addition, according to the actual growth conditions, ripening and coalescence often result in a broad distribution of island sizes. However, potential applications of SiGe nanostructures in novel quantum dot devices require ordered arrangements of the islands, since only a precise definition of their location provides a means to selectively address individual dots. On top of this, size dispersion, i.e. dot-to-dot strain and composition fluctuations, becomes detrimental. As a consequence, a lot of effort has been spent during recent decades to exploit self-organization as a route for the realization of spatially ordered arrays, which in turn is expected to be also a viable path to achieve morphologically ordered ensembles, yielding simultaneously precise control of size, density and alloying of the quantum dots. The aim of this section is to review the experimental attempts and the theory developed in order to understand and control the self-assembly mechanism for Ge quantum dots. Further details can be found in specific reviews on this topic recently published by Montalenti *et al.* (2007), Schmidt (2007), Bauer and Schäffler (2006) and Motta *et al.* (2003).

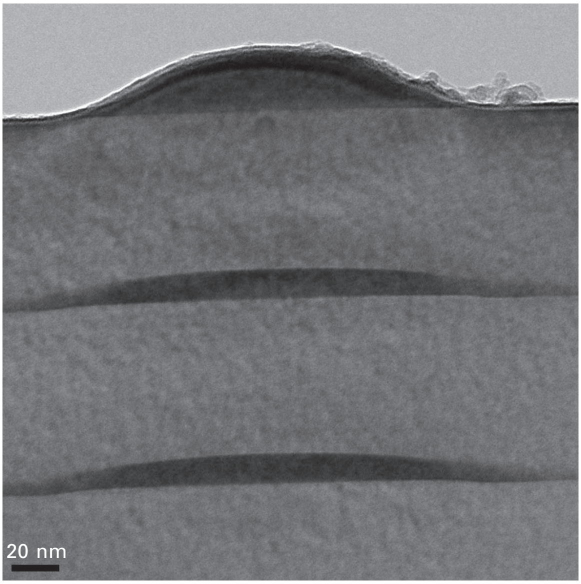
11.3.1 Vertical alignment of Ge dots

A widely used method to achieve a homogeneous size distribution of self-organized dots is to embed several layers of islands within a host matrix. If the thickness of the spacer is chosen appropriately, the anisotropic strain field of the underlying islands induces alignment along the growth direction in a columnar-like fashion (Holý *et al.*, 1999) and results in enhanced island homogeneity (Tersoff *et al.*, 1996; Schittenhelm *et al.*, 1997). Experimentally, the effect of vertical alignment was first observed by Xie *et al.* (1995) and Solomon *et al.* (1996) in III-V based materials. The key physics of the alignment process can be inferred by considering the lattice distortion above a coherent Ge island embedded in Si. Since the self-assembled dot is partially elastically relaxed, the barrier material above the island exhibits a tensile strain field and the lattice host material is widened. For the subsequent island layer, Ge atoms preferably accumulate in those regions of increased lattice parameter and vertical island correlation is induced. An example of ordering is given in the cross-sectional TEM image shown in Fig. 11.7. In order to characterize the vertical ordering of islands, different correlation functions describing a vertical pairing probability have been introduced in the literature, e.g. by Kienzle *et al.* (1999) and Xie *et al.* (1995). It should be noted that although absolute values strongly depend on the specific growth conditions, it can be concluded that, as a general trend, such a probability increases as the spacer thickness is decreased and approaches 100% for spacers tens of nanometers thick.

To a first approximation, neglecting the finite extension of an island, the magnitude of the surface strain field caused by the buried island can be expressed by (Maradudin and Wallis, 1980; Tersoff *et al.*, 1996):

$$\varepsilon(x) = C\,(x^2 + t^2)^{-3/2}\left(1 - \frac{3t^2}{x^2 + t^2}\right) \qquad 11.1$$

Where t is the spacer thickness, x is the lateral direction and C is a constant depending on strain and the volume of the underlying island. The buried island can be modeled as a point-like elastic dipole, which leads to a strain



11.7 Cross-sectional TEM image of vertically aligned Ge islands grown by means of MBE. The sample consists of 8.5 ML Ge separated by 70 nm-thick Si spacers grown at 700°C.

modulation, $\varepsilon(x)$, at the growth front. The strain field along the direction x causes a gradient of the chemical potential and hence atoms preferably nucleate on top of the underlying island. As pointed out by Tersoff *et al.* (1996) two main effects cause the lateral ordering and enhanced size homogeneity: first, two neighboring islands might produce only one larger lattice distortion for the next island layer in case they are too close to each other, and hence merge in the subsequent layer into only one island. Second, the effect of the smaller island is washed out by increasing the number of layers, because of the weakness of their strain field. The concepts and mechanisms pointed out by the aforementioned model stimulated new approaches towards strain engineering at the Si surfaces: by considering a proper shape of the buried island, unexpected patterning on the surface can occur, leading to novel self-organization phenomena (Ni *et al.*, 2004). Marchetti *et al.* (2005) proposed a very simple model to predict the arrangement for the nucleation of Ge islands in a potential well induced by buried truncated-pyramids. Molecular dynamic calculations based on Tersoff interatomic potentials were applied for structures with one-tenth of the actual island dimensions and the relaxed structures were obtained with a 50 ps simulated annealing at 400 K. In the strong elastic interaction limit (Si capping thickness of about 25 nm) a local energy maximum is observed on top of the island center. This feature is absent in the weak limit, i.e. Si capping thickness of about 54 nm. The local maximum is observed for flat-top buried islands. Good agreement between the theoretical predictions and experiments was revealed by atomic force microscopy (AFM) measurements of multilayered SiGe nanostructures grown by chemical vapor deposition.

An additional dot size increase in the upper layers of a stack has been observed in various publications: Kienzle *et al.* (1999), Xie *et al.* (1995), Tersoff *et al.* (1996), Zundel *et al.* (1997), Darhuber *et al.* (1997) and Teichert *et al.* (1996). In particular, upon the deposition of a fixed amount of Ge, Usami *et al.* (2000) found the island density in a second layer to be about four times larger than that in the first layer, along with a transition from dome to pyramid shapes. Although likely mechanisms, such as diffusion of surface atoms (Ledentsov *et al.*, 1996) or merging of laterally closely lying islands (Mateeva *et al.*, 1997; Zundel *et al.*, 1997) take place during growth, the dot size increase strongly suggests that the critical thickness for planar growth in stacked layers is reduced and excess material due to the thinner WL accumulates to form larger islands (Nakata *et al.*, 1997).

In multilayers the modification of the Stranski–Krastanow growth mode, consisting of the reduction of the critical thickness for island nucleation in all but the initial island layer, was clarified by Schmidt *et al.* (1999) by using five-fold stacks of 6.5 ML of Ge, where the Si spacer layer thickness was systematically varied between 10 and 50 nm. Schmidt *et al.* (1999) observed a line-splitting, ΔE_{WL}, of the WL photoluminescence peak and a pronounced blue-shift for the island-related PL peak. Since ΔE_{WL} strongly depends on the elastic interaction between the subsequent layers, it provides a sensitive probe to investigate the strain fields originating from the buried islands. Based on the comparison of the experimental data and finite element calculations, SiGe intermixing taking place during Si overgrowth was addressed, suggesting that SiGe alloying is the responsible effect causing the blue-shift for the PL emission of the islands. By increasing the number of periods, the average lateral island size increases, leading eventually to dislocation nucleation. Such an effect can be avoided by reducing the amount of Ge deposited in the subsequent layers, as can be concluded by *in situ* reflection high-energy electron diffraction (RHEED) investigations, performed by Denker *et al.* (2003) during MBE growth of stacked hut cluster islands. The intensity changes of RHEED spots during the deposition of island stacks enabled Denker *et al.* (2003) to gather inside the stress field induced by buried islands. A numerical calculation of the stress due to truncated pyramids was shown to be in good agreement with experimental data, confirming that Si intermixing in stacked layers plays an important role in defining the onset of SK growth.

11.3.2 Lateral ordering

In view of the integration of self-assembled SiGe islands in novel nanoscale-device architectures (Schmidt and Eberl, 2001b), the exact control of island position on a substrate is mandatory to assure external addressability of the QDs. A bottom-up approach to island ordering on a planar Si surface was first suggested by Shchukin *et al.* (1995). A stable period array of equal-sized islands was expected to be formed because of the elastic strain field interaction between adjacent QDs. A self-ordered array of hut clusters on a square mesh with increasing areal coverage was indeed reported by Floro *et al.* (1998b). The elastic interaction between the strain fields originating from hut clusters or dome clusters has also been used by Tersoff *et al.* (1996) and Mateeva *et al.* (1997) to fabricate self-organized ordered arrays of Ge nanostructures embedded in a Si matrix. An alternative approach based on self-organization at surface steps of vicinal surfaces has been exploited by several authors (see, for instance, Teichert *et al.*, 1998 and Zhu *et al.*, 1998). As an example, controlled positioning of low temperature Ge dots ($T \approx 450°C$) was demonstrated using step bunches on a Si(111) surface by Sgarlata *et al.* (2003). Similarly, lateral ordering has been achieved by means of a strain pattern induced by buried misfit dislocations. This was shown by depositing Ge islands on low-misfit SiGe/Si graded buffer layers (see, for instance, Shiryaev *et al.*, 1997; Kim *et al.*, 2003; Shaleev *et al.*, 2005). In this case, the island pattern turns out to be a replica of the underlying dislocation network, commonly called a cross-hatch pattern (Albrecht *et al.*, 1995). This approach, however, suffers from the detrimental presence of threading dislocations (Ismail *et al.*, 1991), which are known to provide preferential paths for impurity diffusion and to be very efficient recombination centers for minority carriers. Above all, all the aforementioned techniques rely on a self-ordering mechanism, which lacks long-range ordering because it results from the nearest-neighbor and therefore short-range behavior associated with the elastic interaction (Schmidt *et al.*, 2002c).

Since it was soon recognized that suitable engineering of the deposition process was not enough to improve the ordering range of the islands, different concepts based on a proper top-down functionalization of the substrate surface were attempted with different degrees of success. Preferential island nucleation has been shown to occur at the strain modulation created via deposition on a surface on which trenches have been etched (Chen *et al.*, 2008; Tsybeskov *et al.*, 1998; Schmidt *et al.*, 2000b). Very recently Sanduijav *et al.* (2009) performed a detailed STM investigation of the surface evolution with increasing Ge coverage taking place on U- and V- shaped Si stripes (see also Matei *et al.*, 2009). They highlighted how differences in stripe geometry, obtained by a proper tuning of the Si buffer growth parameters, might affect island evolution as well as island positioning. A control of the island nucleation sites has also been obtained by Kamins and Williams (1997), who reported pyramid and dome formation at the edges of Si pillars defined by silicon dioxide windows thermally grown by local oxidation. Although a complete, deterministic control over the formation centers on a Si surface was still a challenge, those attempts elucidated the great potential of island positioning via surface nanotemplates.

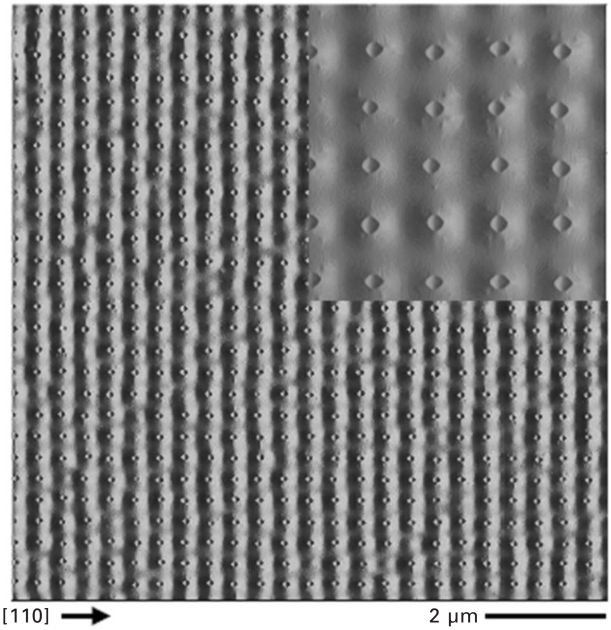
Rapidly, various hybrid approaches combining top-down predefined nucleation sites with bottom-up self-assembly, have been developed and impressive results have already been reported (e.g. Schmidt, 2007). In this process, lithographically defined pits act as a sink for the deposited adatoms, thus allowing the exact positioning of the islands. Besides the possibility of defining ordering on scales limited only by the accuracy and reproducibility of the lithographic technique, growth on prepatterned substrates offers the important advantage of improved island size/shape homogeneity with respect to growth on planar Si(001) substrates. In fact, on planar substrates different islands capture different amounts of the deposited Ge and the ensemble is generally unstable against coarsening. On the other hand, when the island position is predefined by the pattern, the amount of Ge captured by each island mainly depends on the pattern spacing. Moreover, coarsening has been found to be strongly suppressed, allowing the fabrication of homogeneous ensembles of islands with any of the known equilibrium shapes (J.J. Zhang *et al.*, 2007).

Kim *et al.* (1998) demonstrated the possibility of achieving bidimensional arrays with good control over the size and position of the Ge dots on Si substrates. Electron beam lithography was employed to define windows in the oxide ranging from 580 to 90 nm in diameter. Similar results were obtained by Vescan *et al.* (2000) by using optical lithography. In another attempt, Chen *et al.* (2009) employed very recently an intriguing technique to achieve lateral ordering by using nanosphere lithography. The main advantage of this scalable approach relies on its low cost and high throughput. Pit patterns with a periodicity in the range between 100 nm to 1.6 μm were demonstrated by making use of Au-oxide masks realized by self-assembling a monolayer of polystyrene spheres. In this context, it is worth mentioning the investigation of Katsaros *et al.* (2008), who introduced an elegant nanopatterning approach to control the island position by taking advantage of gentle and smooth valleys (about 50 nm wide and 0.5 nm deep) naturally formed on the sample surface during growth: the island trenches. These grooves can be interpreted as footprints left by islands which formed, grew, then shrunk and partially disappeared due to coarsening taking place on the Si surface during either growth or annealing (Rastelli *et al.*, 2005).

Energetic considerations of an empty trench in equilibrium with an island were corroborated by the evidence of lateral motion of the shrinking islands onto their own trenches, leading to the idea of grooves as tiny templates for the subsequent growth of nanostructures. This concept was exploited upon a second-stage deposition of 6 ML Ge at 620°C on a sample where 6 ML Ge were previously grown at 840°C and followed by 300 s of annealing at the same temperature. Katsaros *et al.* (2008) showed that during the second deposition, Ge adatoms form islands that are densely packed along the edges of the trenches. Based on those findings, STM is therefore suggested as a viable approach to exploit deterministic nanoscale pattering (Salling and Lagally, 1994; Kohmoto *et al.*, 1999). An accurate and complete control over island position has been also obtained by patterning the surface with a focused ion beam (Kammler *et al.*, 2003; Hull *et al.*, 2007). Recently, Portavoce *et al.* (2007) described in detail the nucleation mechanism taking place on Ga implanted regions. It was demonstrated how different types of patterns (at least four were observed) can be obtained via tuning the implantation dose and annealing conditions of the substrates. Furthermore, it was clarified how, by a proper choice of the deposition parameters, high homogeneity dot arrays can be achieved. Nevertheless, the authors could not exclude the presence of residual Ga atoms during the Ge growth which might spoil optical and electric properties of the islands.

Nowadays, for the patterning of the Si substrates various types of lithography are routinely used: optical lithography, nanoimprint lithography, e-beam lithography. The pattern transfer into the Si substrates is then usually accomplished by reactive ion etching, leading to regular arrays of nanoholes into the Si substrates (Zhong *et al.*, 2003; Zhong and Bauer, 2004). The islands nucleate preferentially at the bottom of the etched pits, which assume the shape of truncated inverted pyramids, after overgrowth with a thin Si buffer layer. Such spontaneous growth of SiGe QDs in the pits is energetically assisted by the growth kinetics (Zhong *et al.*, 2008). A representative surface morphology is shown in Fig. 11.8. The islands order into a periodic square array, demonstrating that long range ordering of SiGe islands can be achieved. The islands have an average height and diameter of about 5.5 and 108 nm, respectively. A high-resolution scan (2 μm × 2 μm, see inset of Fig. 11.8) shows that, in this case, the islands consist of pyramids with edges aligned along the <100> directions. This nucleation site is at first sight rather surprising, because it appears to be the least favorable site for strain relaxation within the pit template.

To gain a better understanding of the mechanisms that lead to the preferential nucleation of the Ge dots, Chen *et al.* (2006) studied the very early stages of Ge coverage in detail. The experiments showed that the initially forming Ge wetting layer develops a complex, but highly symmetric morphology on the inclined sidewalls of the pits. This pattern is driven by strain and surface energy minimization, and leads after the deposition of typically three monolayers of Ge to a conversion of the pit sidewalls into a pattern that consists exclusively of {105} and {001} facets. Subsequent island nucleation is attributed to Ge accumulation at the bottom of the pits, which is driven by capillarity and the enhanced surface diffusion on the by now {105} faceted



11.8 AFM phase image of laterally ordered SiGe/Si(001) islands. A high-resolution scan (2 μm × 2 μm) is shown in the inset.

sidewalls of the pits, as pointed out theoretically by Biasiol *et al.* (2002). This complex interplay of step bunching and step meandering mechanisms in the confined geometry of the pits leads to preferential dot nucleation at the bottom of the pits.

More recently, Vastola *et al.* (2008) carried out FEM calculations within the framework of anisotropic elasticity theory. Those findings suggested that the mechanisms leading to the island nucleation within pits are driven by specific interplay of different phenomena. In fact, thermodynamic channels have been shown to play an important role, since pit filling provides a better elastic relaxation with respect to mere WL thickening. Such prediction emphasizes that alloying and strain relief pathways for islands grown on a surface with an extrinsic morphology might differ from the ones on planar surfaces. Zhong and Bauer (2004) and Kar *et al.* (2004) were among the first to show that Ge islands have indeed larger volumes and aspect ratios on patterned areas than on corresponding flat surfaces. These observations rather suggest a different compositional and strain state that was recently confirmed via X-ray measurements by Schülli *et al.* (2009) and by a combined approach based on AFM and selective wet chemical etching by Pezzoli *et al.* (2009). Although the mechanisms leading to the composition variations and the morphological evolution of the islands grown on pits are still under debate, it is clear that prepatterning represents a powerful tool to manipulate the strain and composition of the islands. Annealing of islands grown on pit pattern substrates has also been recently shown by J. J. Zhang *et al.* (2009) as a suitable method to control the island morphology via post-growth processing, but without affecting the overall ensemble homogeneity. This uniformity, however, critically depends on the actual size of the pattern area. Kar *et al.* (2004) elucidated that a material depletion region around the finite sized patterned area occurs. The presence of this depletion region has been shown to be due to directional diffusion from the unpatterned, flat surface towards the patterned area and is also reported in other material systems such as InAs/GaAs quantum dots (Kiravittaya *et al.*, 2005). As a consequence, more Ge adatoms accumulate near the inner edge than in the middle of the pattern (Kar *et al.*, 2004), leading to a smooth transition in the island volumes along the pattern direction.

Perfectly ordered SiGe and Ge islands in two (2D) and three (3D) dimensions by providing preferential nucleation sites through nanostructuring for the SK growth mode has been recently realized also by Grützmacher *et al.* (2007). Here they employed extreme ultraviolet interference lithography (EUV-IL) at a wavelength of 13.5 nm for fast, large-area exposure of templates with small periodicity (Dais *et al.*, 2007). Si(001) substrates have been patterned with two-dimensional hole arrays using EUV-IL and reactive ion etching. The gratings used for multiple beam diffraction were fabricated using e-beam lithography from Si substrates coated with a SiN_x and a Cr film. After fabricating the gratings into the Cr film the Si underneath was removed, thus the Cr gratings sit on free standing SiN_x membranes. The depth of the pattern was restricted to 12–14 nm. The pitch width was driven down to tens of nanometers, enabling the method for high-density integration. On these substrates, three-dimensionally ordered SiGe quantum dot crystals with the smallest quantum dot sizes and periods in both the lateral and vertical directions so far have been grown by molecular beam epitaxy (Dais *et al.*, 2008a). X-ray diffractometry from a sample volume corresponding to about 3.6×10^7 dots and atomic force microscopy (AFM) revealed a good structural perfection of the quantum dot crystal and a narrow quantum dot size distribution (Dais *et al.*, 2009). Interband photoluminescence has been observed up to room temperature, indicating a low defect density (Dais *et al.*, 2008b; Grützmacher *et al.*, 2007). Using the Ge concentration and dot shapes determined by X-ray and AFM measurements as input parameters for 3D band structure calculations, a quantitative agreement between measured and calculated PL energies was obtained (Fromherz *et al.*, 2009; Grützmacher *et al.*, 2007). The calculations showed that the band structure of the 3D ordered quantum dot crystal is significantly modified by the artificial periodicity. A calculation of the variation of the eigenenergies based on the statistical variation of the dot dimensions as determined experimentally shows that the calculated electronic coupling between neighboring dots is not destroyed due to the quantum dot size variations (Grützmacher *et al.*, 2007; Holý *et al.*, 2009).

11.4 Rolled-up nanotechnology

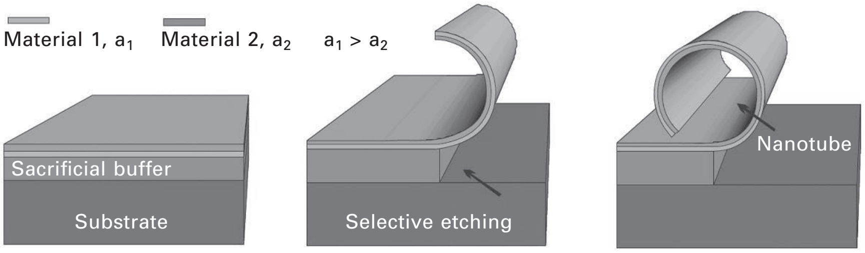
While in the previous sections we explored a bottom-up strategy for nanoscale engineering during growth, here we discuss a combined top-down and bottom-up approach for the 3D fabrication of self-positioned Ge/Si micro- and nano-machined structures based on a stress-induced actuation principle (Prinz *et al.*, 2000; Schmidt and Eberl, 2001a). As outlined above, the germanium lattice constant is larger than that of silicon, i.e. the atoms in Ge are wider spaced than those in Si. If two films of such different crystalline materials are layered to form a thin bilayer in which the atoms across the interface are mutually aligned, then Ge that prefers a larger interatomic spacing is placed under compression, whereas Si is under tension. This built-in strain can be spontaneously relaxed if the sheet is freed from the substrate. As a result, the pair of lattice-mismatched epitaxial layers rolls or curls away, leading to the formation of tubes and coils, respectively (Golod *et al.*, 2001). Nowadays, the knowledge of the atomic structures of the materials can be used in order to deliberately engineer stresses to generate self-folding along preferred directions (Vorob'ev and Prinz, 2002), finally leading to the control of the position as well as the shapes of 3D structures (Schmidt *et al.*, 2001; L.

Zhang *et al*., 2006). These can be tubes (Prinz *et al*., 2001), rings (Schmidt *et al*., 2002a), belts (L. Zhang *et al*., 2005, 2006) and even pop-up systems, recently called micro-origami, since this self-organized method turns out to be similar to the famous Japanese paper folding technique (Vaccaro *et al*., 2001, 2003; Ocampo *et al*., 2003).

Figure 11.9 shows a schematic illustration of a typical roll-up process. The thin, highly strained bilayer rolls up into a micro- or nanotube as soon as the bilayer is detached from the surface by selectively etching the sacrificial buffer layer lying below (Songmuang *et al*., 2006). Since the thin epitaxial bilayers can be grown pseudomorphically onto a semiconductor substrate, the diameter of the tubes, D, can be described by a continuum mechanical model that is linear in single and total layer thickness and inversely proportional to the built-in strain (Nikishkov, 2003; Grundmann, 2003):

$$D = \frac{d_1^4 + 4\chi d_1^3 d_2 + 6\chi d_1^2 d_2^2 + 4\chi d_1 d_2^3 + \chi^2 d_2^4}{3\varepsilon\chi(1+\nu)d_1 d_2(d_1 + d_2)} \qquad 11.2$$

where, ε is the biaxial strain between the two layers, χ the ratio of Young's moduli Y_1/Y_2, ν the Poisson ratio and d_1, d_2 the thicknesses of the first and second layer, respectively. To start the roll-up process the selective etchant needs to have access to the sacrificial layers. The necessary trenches can be placed in the substrate before the strained layer deposition or after the structure is completed. In both cases the patterning can be carried out with lithographic techniques, such as optical or electron beam, allowing individual rolled-up tubes to be deterministically positioned on a substrate surface (Golod *et al*., 2001; Deneke and Schmidt, 2004). After preparation of the trenches, or deposition onto them, the samples are directly inserted into the selective etchant, generally wet chemical solutions. As a consequence, the number of windings, i.e. the number of periods, of the tube wall superlattice can be well defined by a proper definition of the underetching time. It should be noted that this roll-up method is universal and tubes can be fabricated out of many different material systems (see, for instance, Songmuang *et al*.,



11.9 Schematic illustration of the roll-up process to create nanotubes on a substrate surface.

2007a; Schmidt *et al.*, 2002a; Cavallo *et al.*, 2008a). If the tube is made out of rolled-up hybrid layers, such as Si(Ge)/SiO_x which perform multiple rotations, a new type of radial hybrid superlattice (RSL) is created across the wall of the tube (Deneke *et al.*, 2004, 2009; Huang *et al.*, 2005; Meyer *et al.*, 2007). Songmuang *et al.* (2007b) reported an example of a Si/SiO_x RSL. In this case, the necessary strain gradient inside an 8 nm-thick Si layer was obtained by MBE growth on a relaxed Ge buffer deposited on Si. Afterwards, about 7 nm of SiO_x were overgrown by means of thermal evaporation. The resulting Si/SiO_x bilayer was then released from the substrate by wet chemical etching of the Ge sacrificial buffer in hot H_2O_2.

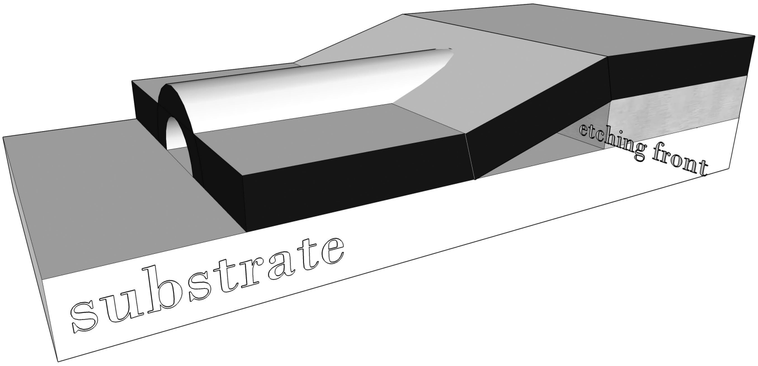
To achieve high selectivities of the etching process, highly boron doped SiGe layers (dopant concentration in the range 10^{20}–10^{21} cm^{-3}) can be grown on undoped Si buffer layers on Si(001) wafers. The resulting structure as shown by TEM cross-section analysis consisted of a radial superlattice with eight windings comprising eight periods of alternating layers of Si and SiO_x. These rolled-up nano-/microtubes are ideal candidates to study novel phenomena of hybrid heterostructures and superlattices both along the tube axis as well as radially across the tube wall superlattice. In particular Si/SiO_x tubes were already exploited as optical resonators, as well as refractometers, and the corresponding modes were found to depend sensitively on the structural uniformity (Songmuang *et al.*, 2007b; Bernardi *et al.*, 2008).

More recently, a second way to achieve strain-driven architectures within the framework of silicon-based systems has been introduced (Cavallo *et al.*, 2007). This approach relies on the silicon on insulator (SOI) technology. It is well known that the aggressive scaling down of Si technology is approaching physical and economic limits, stimulating new and alternative solutions. SOI wafers are indeed already employed in microprocessor production, since the reduced capacitance from the buried oxide improves the circuit performance by up to 30% compared to a bulk-silicon substrate (Paul, 2004). Nevertheless, as power supply voltages continue to decrease with successive scaling, enhancing carrier mobility using strain has become a viable option to sustain a continual drive current increase without traditional scaling. Biaxially tensile-stressed Si on relaxed SiGe on SOI (SGOI) is an ultimate and natural extension of SOI, combining the advantages of SOI and the carrier mobility enhancement of tensile-strained Si for high-performance applications.

Recently, Tezuka *et al.* (2001) proposed a unique and very attractive method to realize SGOI: Ge condensation (Mizuno *et al.*, 2002; Sugii *et al.*, 2002). Typically a diluted SiGe alloy of about 15% of Ge molar fraction is grown on an SOI. Then, the thermal oxidation is performed at a lower temperature than the melting point of SiGe. When SiGe layers are oxidized, only Si atoms are consumed to form SiO_2 whereas Ge atoms are rejected from the oxide film, being bounded between the top and bottom SiO_2 layers. The Si and SiGe layers are merged together, resulting in a relaxed SiGe layer with higher Ge content on top of the buried oxide. SGOI wafers with various SiGe layer thicknesses and Ge profiles can be obtained using different designed oxidation processes. The top oxide as well as the buried oxide, which acts here as the sacrificial layer, can be finally selectively etched in concentrated HF solution to release the semiconductor layer from the substrate. Alternatively, Si layers can also be overgrown on this SGOI substrate prior to etching. The strain profile, the layer composition and the thickness of SiGe on the insulator can be accurately controlled by Ge condensation (Cavallo *et al.*, 2007). The diameter of the rolled-up tubes can therefore be defined by the duration of dry thermal oxidation of the epitaxially grown SiGe layers. This strain-induced technique is rather flexible, allowing tube diameters to be tuned from about 20 nm to the micrometer range (Deneke *et al.*, 2002; Cavallo *et al.*, 2007). Finally, the thickness of the tube wall and the number of contributing layers is directly given by the number of rotations performed during the etching. Consequently, this parameter can be varied by etching time and by definition of the sample area to be rolled by means of lithography (Deneke and Schmidt, 2004).

Upon etching of the sacrificial buffer layer the upper semiconductor layer is released, resulting in an overall relaxation of the film. But, while the internal strains are released, the film might bond back to the substrate, thereby producing wrinkles in the top surface (Bowden *et al.*, 1998; Cerda *et al.*, 2002; Yoo and Lee, 2003). The wrinkling regime has been observed when a uniform or small strain gradient is present in the layer system. An energetic comparison between rolling and wrinkling of compressively strained free-hanging films has been performed by Cendula *et al.* (2009), leading to a phase diagram for the preferential shape of the film as a function of length, average strain, and strain gradient. Very recently, it has been shown that even complex nanochannel networks can be formed on substrate surfaces by deterministic wrinkling (Mei *et al.*, 2007a). This technology is based on the partial release and bond-back of layers and was therefore called REBOLA (Mei *et al.*, 2007a, 2007b).

The evolution of wrinkling can be understood according to the following picture. As schematically shown in Fig. 11.10, a strained film, such as a SiGe layer of a SGOI wafer released through etching from the underlying SiO_2 substrate, naturally forms, by wrinkling, several channels all perpendicular to the front edge that impinge on the Si substrate surface. The wrinkles have high crystalline quality, as confirmed by X-ray analysis, and maintain a small residual stress due to the bond-back effect (Mei *et al.*, 2007b; Malachias *et al.*, 2008). It is worth noting that those channels turn to be interconnected via the area nearby the etching front, i.e. the void region parallel to the unetched SiO_2 under the bending point of the SiGe upper layer. This unique advantage allows the achievement of novel and perfectly controlled high-quality nanotubular building-block structures, paving the way toward the



11.10 Schematic diagram of a structure with individual wrinkling channels.

future realization of nanofluidic devices (Schmidt and Jin-Phillipp, 2001; Mei *et al.*, 2007a; Huh *et al.*, 2007). Wrinkles provide indeed a possible pathway to producing long-range ordered nanochannel networks starting from flat-strained semiconductor films on rigid substrates by using simple and conventional photolithography approaches (Mei *et al.*, 2007a; Malachias *et al.*, 2008). Branched nanochannel networks comprising self-formed wrinkles can be applied to many different material systems, geometries, and layer thicknesses. This topic is of great interest for a number of applications in the field of lab-on-chip-technology and has been recently reviewed by Mei *et al.* (2010).

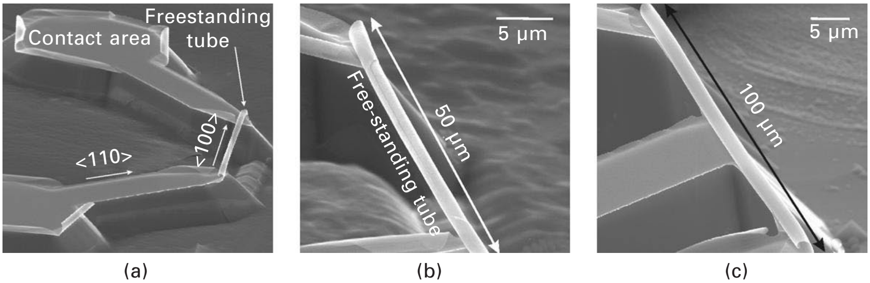
11.5 Potential applications

11.5.1 Electronic transport and spin properties

Much of the progress in the field of semiconductors has been motivated by the need for devices operating at higher frequencies and with lower power dissipation. As semiconductor devices are scaled to smaller dimensions, denser, faster and cheaper integrated circuits are becoming available. Moore's law states that the number of transistors per integrated circuit doubles every 24 months. With minimum feature sizes smaller than 100 nm, Si technology reached the status of a true nanotechnology back in 2001. This enormous achievement is based on a top-down approach that exploits the metal-insulator semiconductor field-effect transistor (MOSFET) scaling rules, pushing the existing devices to their fundamental physical and reliability limits. Moreover, there are application fields such as optoelectronics and spintronics, which are not easily accessible even with further downscaling of available Si devices. In fact, some Si shortcomings such as the low mobility and the indirect bandgap have forced the exploitation of new materials and device concepts. It has been known for many years that by adjusting the semiconductor bandgap and strain, the device switching speed could be boosted. This motivated the introduction of SiGe/Si heterostructures into mainstream Si technology.

Among other concepts, it is worth mentioning the recent realization of n-MOSFETs based on a tensile strained Si n-channel, where the strain is produced by a buried SiGe nanostructure (Donaton *et al.*, 2006; Ang *et al.*, 2007, 2008). In such structures drive current improvements up to 40% and electron mobility enhancements of up to 80% were reported. Moreover, by using SiGe material embedded in Si as p-channel, significant enhancement of the hole mobility can be expected (see, for instance, Myronov *et al.*, 2007), which is important for the realization of complementary MOS (CMOS) circuits. In the approaches proposed by Donaton *et al.* (2006) and Ang *et al.* (2007), the SiGe stressor was obtained either by selective growth (Donaton *et al.*, 2006) or by etching of a planar SiGe layer (Ang *et al.*, 2007). An alternative route is represented by the epitaxial growth of 3D SiGe dots. The possibility of using such islands for faster CMOS was first envisioned by Schmidt and Eberl (2001b) (see also Schmidt *et al.*, 2002b), where the concept of a SiGe dot-based FET (dotFET) was proposed. As shown before, islands with different composition and size can be obtained by proper tuning of the growth and overgrowth parameters, and their position on the substrate can be accurately controlled by performing the Ge growth on Si substrates containing lithographically defined nanoholes (J.J. Zhang *et al.*, 2007; Schmidt, 2007). With respect to epitaxial SiGe layers grown on Si, 3D islands can have substantially larger Ge content, which can possibly create larger strain on the cap layer for n-channel devices and have improved hole mobilities for p-channel devices. Moreover the island lateral size is compatible with next-generation MOSFETs, which makes this system very attractive.

Three-dimensional micro- and nanoarchitectures fabricated by the deterministic release and rearrangement of nanomembranes on substrate surfaces might also provide a new route to investigate novel electron transport properties (see, for instance, Friedland *et al.* (2009) and references therein). Conductance in cylindrical structures is expected to be highly sensitive to the curvature because of the confinement potentials induced by the tube geometry (Meyer *et al.*, 2007; Ferrari and Cuoghi, 2008). The peculiar curved structure of such surfaces might allow the investigation of topographical quantum effects, which are beyond the grasp of their planar counterpart (Shaji *et al.*, 2007). A similar approach has been recently put forward by Cavallo *et al.* (2008b), who characterized the electrical response of freestanding highly boron-doped SiGe/Si rolled-up tubes, grown by molecular beam epitaxy on p-type Si(001) wafers. Metal contacts were evaporated on the unreleased layers, which were patterned by optical lithography and reactive ion etching. The subsequent selective etching of the sacrificial undoped buffer layer produced 50 μm or 100 μm long freestanding tubes as shown in Fig. 11.11.



11.11 SEM images of freestanding microtubes including contact pads. (b) and (c) Higher resolution SEM images of 50 μm and 100 μm long highly boron doped $Si_{0.6}Ge_{0.3}$:B/Si:B tubes (courtesy Dr F. Cavallo).

This demonstrates the possibility of separating rolled-up tubes from the substrate, enabling a device geometry, which easily integrates the system with contact pads for lateral transport. This simple ohmic device based on a freestanding SiGe:B/Si:B tube can indeed be regarded as a good candidate for gas sensor (Houlet *et al.*, 2008) and electrically driven micropump (Tsai and Sue, 2007). Finally, the technique of rolling up films can be used to fabricate also helix and ring structures (L. Zhang *et al.*, 2005). These structures are currently explored as functional building blocks for nanorobotics used as artificial bacterial flagella (L. Zhang *et al.*, 2008a, 2008b, 2009).

Besides the interest in the realization of devices with improved performances, SiGe heterostructures are receiving attention for their potential also in the field of spintronics because of the long electron spin coherence time in Si (Appelbaum *et al.*, 2007). Furthermore, Si and SiGe are materials that can be isotopically purified with nuclei of zero spins unlike GaAs-based systems. Kane (1998) proposed a very interesting and elegant design of qubits devices based on nuclear spins of phosphorus dopant atoms embedded in a silicon host. This approach has stimulated novel implementations of spin-resonance transistors in Si, accomplished by means of band-structure engineering in heterostructures. By working with charge spins rather than nuclear spins, the requirement of a sophisticated spin transfer between electrons and nuclei, for read-in-read-out of quantum data and for the operation becomes less demanding and more feasible. In addition, for SiGe nanostructures coherently embedded in Si the band-edge alignment is a type II alignment (Schmidt *et al.*, 2000c). This leads to localized confining potentials for electrons in the Si matrix. As a consequence, self-assembled SiGe islands can assure precise manipulation via an external addressability of the localized electron spin states, which is an indispensable precondition for the realization of spintronic devices (Friesen *et al.*, 2003). However, it is worth noting that this field is still in its infancy and the demonstration of even basic spintronic phenomena such as spin injection and detection has not been completely achieved (Žutić *et al.*, 2004).

Very recently, Katsaros *et al.* (2010) carried out a breakthrough investigation on hole transport through single SiGe nanostructures. In this work, single domes were laterally contacted with Al electrodes by means of a nanolithographic process. Islands were grown on silicon-on-insulator where a highly doped Si substrate was acting as a back-gate. By taking advantage of the geometry of such superconductor–semiconductor devices, Katsaros *et al.* (2010) reported single-hole super-current with discrete spectra features, providing robust evidence (at low temperature) of full 3D quantum confinement within dome-shaped islands. Furthermore, it was shown that SiGe islands can be regarded as good candidates for coherent spin manipulation by means of gate-controlled electric fields. Experimental findings pointed out that strong spin-orbit coupling can be achieved and controlled to some extent via external magnetic fields. Besides, the values of the hole g-factors turned out to be appreciably anisotropic ($g_\perp = 2.71$ and $g_\parallel = 1.21$) and affected by the number of the confined holes, therefore suggesting that g-factors in SiGe nanostructures are defined by the orbital wavefunctions.

11.5.2 Thermal transport

Thermoelectric materials convert thermal gradients and electric fields for power generation and for refrigeration, respectively. Thermoelectric devices currently find only limited applications, because of their poor efficiency, which is benchmarked by the so-called thermoelectric figure of merit:

$$\mathrm{ZT} = \frac{S^2\sigma}{\kappa}T \qquad 11.3$$

which includes the Seebeck coefficient S, the average temperature T, the electrical conductivity σ, and the thermal conductivity κ. The latter contains both electronic and phononic contributions. Maximizing ZT is challenging, because optimizing one physical parameter often adversely affects another. In metals, electrons contribute equally to the electrical and thermal transport (Wiedemann–Franz law); in insulators only phonons give a non-negligible contribution to thermal conductivity; in semiconductors both subsystems strongly contribute to the thermal transport, while the electrical conductivity can be dramatically changed via doping. For semiconductors, it has been historically challenging to obtain a figure of merit higher than 1, because of a bottleneck due to high thermal conductivities. One of the lower values for the thermal conductivity in solids is generally that of an alloy and is often referred to as the ‘alloy limit’.

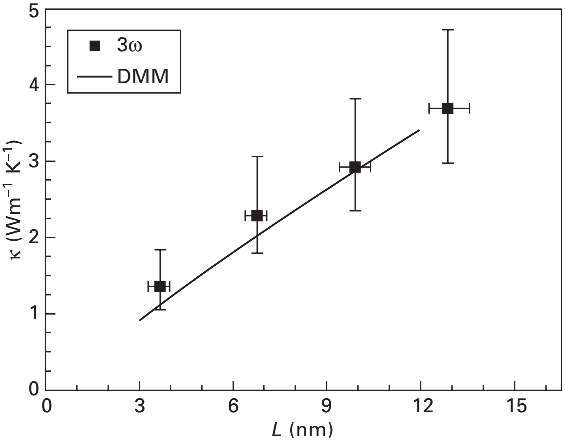
Recently, it was shown that nanostructures provide scattering mechanisms other than alloy scattering. By doing so, researchers were able to demonstrate extremely low thermal conductivities even in semiconductors. For instance, by embedding epitaxial ErAs nanoparticles into an InGaAs alloy matrix, short-wavelength phonons can be scattered by the matrix while mid- and long-wavelength phonons can be scattered by the nanoparticles. This approach leads to thermal conductivity clearly below that of the alloy film, without any detrimental effect on the figure of merit, because of the high crystalline quality of the epitaxial films (Kim *et al.*, 2006). Dimensionality reduction toward the nanoscale provides indeed an additional parameter to control the thermal as well as the electrical response of thermoelectric materials.

Hicks and Dresselhaus (1993a, 1993b) and Hicks *et al.* (1993) were the first to show that ZT could be increased by preparing materials, like Bi_2Te_3, in a quantum-well superlattice structure; they also showed a strong influence of anisotropy on achievable ZT values. Besides, quantum confinement creates sharp features in the electronic density of states and increases the difference between conductivity of hot and cold carriers, finally leading to an increase of the power factor ($S^2\sigma$). Subsequent experimental works on Bi_2Te_3 material systems confirmed high ZT values (Venkatasubramanian *et al.*, 2001; Böttner *et al.*, 2006), but also mentioned the difficulties of electric contacts, materials matching, and fabrication. In this respect, silicon-based systems are expected to improve the overall quality of the devices: the well-known growth processes and the well-established contact formation represent a valuable resource (Shakouri, 2006; Y. Zhang *et al.*, 2006; Ezzahri *et al.*, 2007, 2008). Reducing the layer thickness of a given material is thus expected to lead to new electronic, mechanical and thermodynamic properties, which are distinct from those of bulk material.

P.P. Zhang *et al.* (2006a,b) demonstrated that electronic conduction in thin silicon-on-insulator membranes is determined not by bulk dopants but by the interaction of surface or interface electronic energy levels with the band structure of the thin silicon template layer. To further increase the interface effects, the construction of novel types of heterostructures such as strain-induced architectures with well-defined interfaces in between would be desirable and might extend the concept of quantum-well superlattices. For example Si/SiO_x radial superlattices are promising candidates in this field, since unperturbed charge carrier transport might take place via tunneling through the oxide film that might simultaneously act as a phonon blocking barrier. Similarly, semiconductors and metallic-based superlattices such as Cr/Si(Ge) tubes (Cavallo *et al.*, 2008a) have the potential to achieve thermoelectric power factors significantly larger than bulk values with ZT of the order of 5 or 6, because of the thermoionic emission enhancement (see Shakouri, 2006). Rolled-up technology offers the additional possibilities to nearly eliminate the substrate influence, to provide structures which allow a radial heat transfer, and finally to naturally integrate all necessary contacts for sensing and measurement (temperature determination and adjustment, electrical characterization). However, there is certainly more to be done on the thermoelectric properties of the semiconductor 3D structures from both an experimental and a theoretical point of view.

We would like to stress that the overall improvement of the figure of merit can be obtained also in other low-dimensional systems. For instance, Harman *et al.* (2002) have considered PbSeTe quantum dots (QDs) grown by means of Stranski–Krastanow on PbTe. They investigated a multilayer composed of more than 8000 layers of PbSeTe/PbTe QDs and built up a test device with a claimed ZT of about 2 (see, however, Koh *et al.*, 2009; Vineis *et al.*, 2008). As far as SiGe concerns, a cross-plane thermal conductivity lower than the 'alloy limit' value for the bulk material was first reported by Yang *et al.* (2002a). A room-temperature (RT) value of 2.9 W m^{-1} K^{-1} was obtained by using an *n*-type SiGe/Si QD multilayer comprising ~15 Å of Ge separated by 75 Å-thick Si spacer layers. A slightly higher value of 6.2 W m^{-1} K and an in-plane thermal conductivity of ~30 W m^{-1} K^{-1} were previously reported by Liu *et al.* (2001) by using again ~15 Å of Ge, but separated this time by 200 Å-thick Si spacer layers. Later, Liu *et al.* (2003) addressed the effect of the Ge content of the islands on the thermal conductivity. By growing a set of samples at different substrate temperatures and different Ge coverages, it was demonstrated that thermal conductivity can be reduced either by a low-temperature deposition or by increasing the amount of the deposited Ge. Those findings were ascribed to an enhanced phonon mismatch with the Si matrix, due to the Ge-richer composition of the islands. Lee and Venkatasubramanian (2008) studied in detail the effect of island density and period length on the RT cross-plane thermal conductivity. A monotonic decrease of κ was observed by increasing the areal coverage from 8% up to 20%. More strikingly, for short periods of 3.9 nm, the thermal conductivity reached values in the range 2.0–2.7 W m^{-1} K^{-1}. This decrease of κ was justified by phonon interface scattering. Despite the ongoing debate concerning how the limits of lattice thermal conductivity can be broken and to that extent they can be pushed, nanostructuring seems to provide a straightforward approach to obtaining ultra-low thermal conductivities, opening up new and intriguing possibilities within the framework of the rapidly growing field of thermoelectrics.

So far, the lowest thermal conductivity value for any SiGe-based material fully compatible with current Si nanoelectronics has been reported by Pernot *et al.* (2010). The cross-plane thermal conductivity of SiGe nanodot multilayers grown by means of solid-source MBE on silicon was measured by using two independent state-of-the-art techniques, namely heterodyne picosecond thermoreflectance and the 3-ω method, yielding compatible results. By engineering a set of individual phonon-scattering nanodot barriers, Pernot *et al.* (2010) demonstrated that the thermal conductivity decreases linearly with decrease of the Si interlayer spacing (see Fig. 11.12), providing a means



11.12 Thermal conductivities, κ, for SiGe/Si quantum dot multilayers (comprising 11 layers), measured by the 3ω technique, as a function of the average distance between layers, *L*. The solid line is the result of the Diffuse Mismatch Model (DMM) (Pernot *et al.*, 2010).

to accurately tailor the thermal conductivity to extremely low values: for the smallest superlattice period length (about 3 nm), and 5 Ge layers, the thermal conductivity reaches a value as low as 0.9 W m^{-1} K^{-1}, even lower than the that amorphous Si. The highly diffuse character of the interfaces led indeed to measured thermal resistances associated with one individual interface in the range of 2.5–4 $\times 10^{-9}$ m^2 W^{-1} K^{-1}, in close agreement with the diffuse mismatch model and about three times larger than the data previously reported in the literature (Pernot *et al.*, 2010). Atomistic Green's function calculations elucidated the effect of the nanodots: in the low-frequency regime, the calculated phonon mean free paths for the nanodot multilayers are considerably shorter than those for the flat layer counterparts, whereas the two are comparable at high frequency. This suggests that the dramatic reduction of the thermal conductivity can be linked to the roughness of the larger feature introduced by the dots. Such achievement may be relevant to the development of integrated miniaturized energy harvesting or thermal management devices, in view of the integrability of SiGe in novel nanoscale devices.

It is worth mentioning that, in spite of its importance, there are no reports concerning cross-plane electric conductivity for SiGe/Si QDs. Only in-plane mobility has been measured by a few groups: Bao *et al.* (2005) and Yakimov *et al.* (2006). Besides, to date, Yang *et al.* (2002b) have been the only researchers addressing cross-plane thermopower in SiGe nanostructure-based systems. Their findings suggest only a moderate increase of S with respect to the value of bulk Si. Due to the lack of experimental data, it is still not clear whether the presence of QDs can affect electronic transport through a multilayer of QDs and whether *S* can be further increased by proper design of the QD layers. Nevertheless, the reduction of interlayer spacing can be foreseen as a viable path to put forward the efficiency of SiGe-based thermoelectrics. Indeed, the thickness of the spacers seems to be a key factor in achieving lower values of κ (Lee and Venkatasubramanian, 2008; Pernot *et al.*, 2010) and to be a beneficial parameter for the improvement of the electronic transport due to the formation of superlattice minibands (Wacker, 2002).

11.6 Sources of further information and advice

SiGe micro- and nanostructures have established themselves as a means to enable technologies in different fields ranging from electronics to thermoelectrics. This chapter represents a timely attempt to collect the main information about 3D architectures based on strain engineering within the huge and widely dispersed volume of data describing an important semiconductor such as SiGe. It is therefore fitting to dedicate this last section specifically to the main research groups whose works have made the field successful. The reader should, however, be aware that because of the fast-growing interest in SiGe micro- and nanoarchitectures it is simply impossible to report and describe all the contributors to the field. Nevertheless, further details can be found in the most relevant books and review articles that have been reported in the previous sections along with the discussions about the specific topics they are describing.

11.6.1 Survey of research centers

The IFW Dresden (Germany) is a non-university research institute, and a member of the Leibniz Association. It is concerned with modern materials science and combines explorative research in physics and chemistry with technological development of new systems and products. Its research focuses on functional materials such as superconducting and magnetic materials, thin films and nanostructures, which hold a key position in many applications. The Leibniz Institute for Solid State and Materials Research Dresden is divided into five institutes, one of which is the Institute for Integrative Nanosciences (IIN) headed by Prof. O. G. Schmidt (website: http://www.ifw-dresden.de/institutes/iin). The IIN in turn consists of two departments, one for Quantum Dots (headed by Dr A. Rastelli), and the other for Strained Nanoarchitectures (headed by Prof. O. G. Schmidt). The IIN group has been working on the growth of SiGe dots for more than 10 years and has made a major contribution to the fundamental understanding of the formation and properties of SiGe islands during epitaxial growth of Ge on Si and during their encapsulation in a Si matrix. Optical properties for all types of SiGe islands have been systematically explored and deep understanding on band-edge alignment in single and stacked layers has been acquired. Above all, the group has pioneered the field of integrative rolled-up tubes, establishing cutting-edge technologies and a well-recognized position in the development of fluidic devices, microtube ring resonators for optofluidic sensing, and rolled-up electronics, as well as hyper-lenses for high-resolution local imaging and novel concepts for micro-/nanorobotics.

The Research Centre Jülich with 4500 employees is one of the biggest research institutions in Germany. It is a member of the Helmholtz Association. The research topics span from fundamental research to application-oriented research in the fields of health, environment, matter, energy, and information technology. The Institute for Bio- and Nanotechnology-1 at the Forschungszentrum Jülich GmbH is directed by Prof. D. Grützmacher and provides a platform for the fabrication and development of semiconductor nanodevices in the III/V as well as in the Si-based material system (website: http://www.fz-juelich.de/ibn/ibn1_en/). The group has a large amount of experience in the fabrication and analysis of these semiconductor nanodevices.

The Institute for Semiconductor and Solid State Physics at J.-Kepler University Linz (Austria) has existed for 10 years (web page: http://www.hlphys.jku.at/). It is currently headed by Prof. Dr Friedrich Schäffler. It has facilities for the growth of semiconductor hetero- and nanostructures, for their structural, electronic and optical characterization. In particular, the X-ray group of Prof. Dr G. Bauer has been involved in a number of EC projects, particularly in the area of the structural characterization of Si/SiGe and Si/SiGeC nanostructures and of intersubband transitions in Si/SiGe multiquantum wells suitable for detector applications. The group has successfully collaborated with a number of universities and industrial laboratories. It currently carries out activities in semiconductor physics, concerning growth, structural, electronic and optical investigations of semiconductor hetero- and nanostructures. The Institute is structured in two groups which have facilities for the MBE growth of SiGeC-based heterostructures and of IV-VI heterostructures. It has characterization facilities such as ellipsometry, Auger-electron spectroscopy, atomic force and scanning tunneling microscopy, high-resolution X-ray diffraction and X-ray reflectivity for ex-situ structural analysis. In a clean room, basic fabrication steps (optical and e-beam lithography, etching, deposition of oxides, nitrides, etc.) can be performed.

The University of Milano-Bicocca is a recent and dynamic institution located in Milan (Italy). In particular, the Department of Materials Science gathers most of the expertise in Solid State Physics and Materials Chemistry, formerly present in the University of Milano, its original institution. The Department hosts a well-established internationally renowned experimental and theoretical research activity in the field of inorganic semiconductors and semiconductor nanostructures. In particular, the theory group of Prof.

Leo Miglio (web page: http://www.mater.unimib.it/utenti/migliogroup/) had focused its research activities on the following topics: (i) elastic relaxation of stress and strain engineering at Ge-rich nanostructures on/in silicon; (ii) kinetic and thermodynamic issues in the morphological evolution of Ge-based epitaxial nanostructures on silicon, and (iii) plastic relaxation of stress in SiGe and Ge heterostructures. These activities are supported mainly by calculations based on the finite element method and by empirical Tersoff-potential large-scale molecular dynamics simulations.

At the University of Wisconsin-Madison (USA), the group headed by Prof. M. Lagalli (website: http://mrgcvd.engr.wisc.edu/index.html) focuses on the creation and determination of nanoscale properties of Group IV semiconductors. In particular they are investigating silicon-based nanomembranes both experimentally and theoretically, aiming to determine the mechanisms leading to the formation of novel surface and interface structures and nano-morphologies, and the relationship between such nano-morphologies and localized electronic and optoelectronic properties. Current areas of active research are, among others, (i) uniaxial and biaxial strain in semiconductors; (ii) quantum dot epitaxy on suspended Si membranes; (iii) Ge stressors deposited on free-standing Si membranes; and (iv) Si nanomembranes on plastics.

The research staff at the IBM Thomas J. Watson Research Center (New York City, USA), in particular Dr J. Tersoff (http://www.research.ibm.com/atomic/nano/), is actively involved in the theoretical investigation of the electronic structure and morphological properties of surfaces and interfaces.

ICMAB, Institut de Ciència de Materials de Barcelona (Spain), is a research institute more than two decades old. At present, it includes about 200 people with 53 scientists as permanent staff. The nanostructured material department (http://www.icmab.es/icmab/organ/dep/nanostructured-materials.html), directed by Dr I. Alonso Carmona, is mainly concerned with MBE growth of low-dimensional nanostructures based on Si, Ge and C. The research activity is mainly focused on the control of Ge island nucleation via submonolayer deposition of C on Si surfaces.

The National Research Council, NRC, Institute for Microstructural Sciences in Ottawa, Canada (http://www.nrc-cnrc.gc.ca/eng/ibp/ims.html) undertakes experimental research into the electronic and optical properties of semiconductor-based materials including SiGe quantum dots. Much of the research is directed towards quantum information processing and advanced photonic devices, with special attention to nanostructured materials emitting in the near-infrared region of the spectrum. Much effort is also dedicated to understanding the fundamental properties of these complex systems, the relevant properties of the underlying materials and the mechanisms responsible for their ultimate behavior.

The Institute of Robotics and Intelligent Systems (http://www.iris.ethz.ch/msrl/) at ETH Zürich (Switzerland) pursues a research program that includes works on nanorobotics based on self-rolled and curled SiGe nanostructures. IRIS research develops the tools and processes required to fabricate and assemble micron-sized robots and nanometer-scale robotic components. Many of these systems are used for robotic exploration within biological domains, such as in the investigation of molecular structures, cellular systems, and complex organism behavior.

11.7 Acknowledgments

The authors would like to thank F. Cavallo, T. Dienel, M. Stoffel, A. Rastelli and N. Mingo for fruitful discussions and the material provided for this chapter, and B. Rellinghaus and T. Gemming for the use of TEM. FP would like to thank 'Dote ricercatori': FSE, Regione Lombardia, Italy.

11.8 References

Albrecht, M, Christiansen, S, Michler, J, Dorsch, W, Strunk, H P, Hansson, P O and Bauser, E (1995) 'Surface ripples, crosshatch pattern, and dislocation formation – cooperating mechanisms in lattice mismatch relaxation', *Applied Physics Letters*, 67(9), 1232–1234

Ang, K W, Tung, C H, Balasubramanian, N, Samudra, G S and Yeo, Y C (2007) 'Strained *n*-channel transistors with silicon source and drain regions and embedded silicon/germanium as strain-transfer structure', *IEEE Electron Device Letters*, 28(7), 609–612

Ang, K W, Lin, J, Tung, C H, Balasubramanian, N, Samudra, G S and Yeo, Y C (2008) 'Strained *n*-MOSFET with embedded source/drain stressors and strain-transfer structure (STS) for enhanced transistor performance', *IEEE Transactions on Electron Devices*, 55(3), 850–857

Appelbaum, I, Huang, B Q and Monsma, D J (2007) 'Electronic measurement and control of spin transport in silicon', *Nature*, 447(7142), 295–298

Asaro, R J and Tiller, W A (1972) 'Interface morphology development during stress-corrosion cracking. 1. Via surface diffusion', *Metallurgical Transactions*, 3(7), 1789–1796

Bao, Y, Liu, W L, Shamsa, M, Alim, K, Balandin, A A and Liu, J L (2005) 'Electrical and thermal conductivity of Ge/Si quantum dot superlattices', *Journal of the Electrochemical Society*, 152(6), G432–G435

Baribeau, J M, Wu, X, Rowell, N L and Lockwood, D J (2006) 'Ge dots and nanostructures grown epitaxially on Si', *Journal of Physics – Condensed Matter*, 18(8), R139–R174

Bauer, G and Schäffler, F (2006) 'Self-assembled Si and SiGe nanostructures: New growth concepts and structural analysis', *Physica Status Solidi A – Applications and Materials Science*, 203(14), 3496–3505

Berbezier, I and Ronda, A (2009) 'SiGe nanostructures', *Surface Science Reports*, 64(2), 47–98

Berbezier, I, Ronda, A and Portavoce, A (2002) 'SiGe nanostructures: new insights into growth processes', *Journal of Physics – Condensed Matter*, 14(35), 8283–8331

Berbezier, I, Ronda, A, Volpi, F and Portavoce, A (2003) 'Morphological evolution of SiGe layers', *Surface Science*, 531(3), 231–243

Bernardi, A, Kiravittaya, S, Rastelli, A, Songmuang, R, Thurmer, D J, Benyoucef, M and Schmidt, O G (2008) 'On-chip Si/SiO$_x$ microtube refractometer', *Applied Physics Letters*, 93(9), 094106

Biasiol, G, Gustafsson, A, Leifer, K and Kapon, E (2002) 'Mechanisms of self-ordering in nonplanar epitaxy of semiconductor nanostructures', *Physical Review B*, 65(20), 205306

Böttner, H, Chen, G and Venkatasubramanian, R (2006) 'Aspects of thin-film superlattice thermoelectric materials, devices, and applications', *MRS Bulletin*, 31(3), 211–217

Boukai, A I, Bunimovich, Y, Tahir-Kheli, J, Yu, J K, Goddard, W A and Heath, J R (2008) 'Silicon nanowires as efficient thermoelectric materials', *Nature*, 451(7175), 168–171

Bowden, N, Brittain, S, Evans, A G, Hutchinson, J W and Whitesides, G M (1998) 'Spontaneous formation of ordered structures in thin films of metals supported on an elastomeric polymer', *Nature*, 393(6681), 146–149

Brehm, M, Montalenti, F, Grydlik, M, Vastola, G, Lichtenberger, H, Hrauda, N, Beck, M J, Fromherz, T, Schäffler, F, Miglio, L and Bauer, G (2009) 'Key role of the wetting layer in revealing the hidden path of Ge/Si(001) Stranski–Krastanow growth onset', *Physical Review B*, 80(20), 205321

Brunner, K (2002) 'Si/Ge nanostructures', *Reports on Progress in Physics*, 65(1), 27–72

Capasso, F (1987) 'Band-gap engineering – from physics and materials to new semiconductor-devices', *Science*, 235(4785), 172–176

Capellini, G, De Seta, M, Spinella, C and Evangelisti, F (2003) 'Ordering self-assembled islands without substrate patterning', *Applied Physics Letters*, 82(11), 1772–1774

Capellini, G, De Seta, M, Di Gaspare, L, Evangelisti, F and d'Acapito, F (2005) 'Evolution of Ge/Si(001) islands during Si capping at high temperature', *Journal of Applied Physics*, 98(12), 124901

Cavallo, F, Songmuang, R, Ulrich, C and Schmidt, O G (2007) 'Rolling up SiGe on insulator', *Applied Physics Letters*, 90(19), 193120

Cavallo, F, Sigle, W and Schmidt, O G (2008a) 'Controlled fabrication of Cr/Si and Cr/SiGe tubes tethered to insulator substrates', *Journal of Applied Physics*, 103(11), 116103

Cavallo, F, Songmuang, R and Schmidt, O G (2008b) 'Fabrication and electrical characterization of Si-based rolled-up microtubes', *Applied Physics Letters*, 93(14), 143113

Cendula, P, Kiravittaya, S, Mei, Y F, Deneke, C and Schmidt, O G (2009) 'Bending and wrinkling as competing relaxation pathways for strained free-hanging films', *Physical Review B*, 79(8), 085429

Cerda, E, Ravi-Chandar, K and Mahadevan, L (2002) 'Thin films – Wrinkling of an elastic sheet under tension', *Nature*, 419(6907), 579–580

Chaparro, S A, Zhang, Y, Drucker, J, Chandrasekhar, D and Smith, D J (2000) 'Evolution of Ge/Si(100) islands: Island size and temperature dependence', *Journal of Applied Physics*, 87(5), 2245–2254

Chen, G, Lichtenberger, H, Bauer, G, Jantsch, W and Schäffler, F (2006) 'Initial stage of the two-dimensional to three-dimensional transition of a strained SiGe layer on a pit-patterned Si(001) template', *Physical Review B*, 74(3), 035302

Chen, G, Vastola, G, Lichtenberger, H, Pachinger, D, Bauer, G, Jantsch, W, Schaeffler, F and Miglio, L (2008) 'Ordering of Ge islands on hill-patterned Si(001) templates', *Applied Physics Letters*, 92(11), 113106

Chen, P X, Fan, Y L and Zhong, Z Y (2009) 'The fabrication and application of patterned Si(001) substrates with ordered pits via nanosphere lithography', *Nanotechnology*, 20(9), 095303

Costantini, G, Rastelli, A, Manzano, C, Songmuang, R, Schmidt, O G, Kern, K and von Känel, H (2004) 'Universal shapes of self-organized semiconductor quantum dots: Striking similarities between InAs/GaAs(001) and Ge/Si(001)', *Applied Physics Letters*, 85(23), 5673–5675

Cullis, A G and Booker, G R (1971) 'Epitaxial growth of silicon and germanium films on (111) silicon surfaces using UHV sublimation and evaporation techniques', *Journal of Crystal Growth*, 9(1), 132–138

Dais, C, Solak, H H, Ekinci, Y, Müller, E, Sigg, H and Grützmacher, D (2007) 'Ge quantum dot molecules and crystals: Preparation and properties', *Surface Science*, 601(13), 2787–2791

Dais, C, Solak, H H, Müller, E and Grützmacher, D (2008a) 'Impact of template variations on shape and arrangement of Si/Ge quantum dot arrays', *Applied Physics Letters*, 92(14), 143102

Dais, C, Mussler, G, Sigg, H, Fromherz, T, Auzelyte, V, Solak, H H and Grützmacher, D (2008b) 'Photoluminescence studies of SiGe quantum dot arrays prepared by templated self-assembly', *EPL*, 84(6), 67017

Dais, C, Mussler, G, Sigg, H, Mueller, E, Solak, H H and Grützmacher, D (2009) 'Evolution and stability of ordered SiGe islands grown on patterned Si(100) substrates', *Journal of Applied Physics*, 105(12), 122405

Darhuber, A A, Holý, V, Stangl, J, Bauer, G, Krost, A, Heinrichsdorff, F, Grundmann, M, Bimberg, D, Ustinov, V M, Kopev, P S, Kosogov, A O and Werner, P (1997) 'Lateral and vertical ordering in multilayered self-organized InGaAs quantum dots studied by high resolution X-ray diffraction', *Applied Physics Letters*, 70(8), 955–957

Daruka, I and Tersoff, J (2002) 'Existence of shallow facets at the base of strained epitaxial islands', *Physical Review B*, 66(13), 132104

Daruka, I, Tersoff, J and Barabasi, A L (1999) 'Shape transition in growth of strained islands', *Physical Review Letters*, 82(13), 2753–2756

De Seta, M, Capellini, G, Evangelisti, F and Spinella, C (2002) 'Intermixing-promoted scaling of Ge/Si(100) island sizes', *Journal of Applied Physics*, 92(1), 614–619

De Seta, M, Capellini, G, Di Gaspare, L, Evangelisti, F and D'Acapito, F (2006) 'Freezing shape and composition of Ge/Si(001) self-assembled islands during silicon capping', *Journal of Applied Physics*, 100(9), 093516

Deneke, C and Schmidt, O G (2004) 'Real-time formation, accurate positioning, and fluid filling of single rolled-up nanotubes', *Applied Physics Letters*, 85(14), 2914–2916

Deneke, C, Müller, C, Jin-Phillipp, N Y and Schmidt, O G (2002) 'Diameter scalability of rolled-up In(Ga)As/GaAs nanotubes', *Semiconductor Science and Technology*, 17(12), 1278–1281

Deneke, C, Jin-Phillipp, N Y, Loa, I and Schmidt, O G (2004) 'Radial superlattices and single nanoreactors', *Applied Physics Letters*, 84(22), 4475–4477

Deneke, C, Schumann, J, Engelhard, R, Thomas, J, Sigle, W, Zschieschang, U, Klauk, H, Chuvilin, A and Schmidt, O G (2008) 'Fabrication of radial superlattices based on different hybrid materials', in Hirayama, Y and Sogawa, T, eds, *Physica Status Solidi C – Current Topics in Solid State Physics*, 5(9), 2704–2708

Deneke, C, Songmuang, R, Jin-Phillipp, N Y and Schmidt, O G (2009) 'The structure of hybrid radial superlattices', *Journal of Physics D: Applied Physics*, 103001 (16 pp)

Denker, U, Schmidt, O G, Jin-Philipp, N Y and Eberl, K (2001) 'Trench formation around and between self-assembled Ge islands on Si', *Applied Physics Letters*, 78(23), 3723–3725

Denker, U, Stoffel, M and Schmidt, O G (2003) 'Quantitative evaluation of stress-field attenuation in stacks of self-assembled Ge islands', *Applied Physics Letters*, 83(7), 1432–1434

Donaton, R A, Chidambarrao, D, Johnson, J, Chang, P, Liu, Y C, Henson, W K, Holt, J, Li, X, Li, J, Domenicucci, A, Madan, A, Rim, K and Wann, C (2006) 'Design and fabrication of MOSFETs with a reverse embedded SiGe (Rev e-SiGe) structure', in *2006 International Electron Devices Meeting*, Vols 1 and 2, 191–194

Drucker, J and Chaparro, S (1997) 'Diffusional narrowing of Ge on Si(100) coherent island quantum dot size distributions', *Applied Physics Letters*, 71(5), 614–616

Eaglesham, D J and Cerullo, M (1990) 'Dislocation-free Stranski–Krastanow growth of Ge on Si(100)', *Physical Review Letters*, 64(16), 1943–1946

Ezzahri, Y, Grauby, S, Rampnoux, J M, Michel, H, Pernot, G, Claeys, W, Dilhaire, S, Rossignol, C, Zeng, G and Shakouri, A (2007) 'Coherent phonons in Si/SiGe superlattices', *Physical Review B*, 75(19), 195309

Ezzahri, Y, Zeng, G, Fukutani, K, Bian, Z and Shakouri, A (2008) 'A comparison of thin film microrefrigerators based on Si/SiGe superlattice and bulk SiGe', *Microelectronics Journal*, 39(7), 981–991

Ferrari, G and Cuoghi, G (2008) 'Schrodinger equation for a particle on a curved surface in an electric and magnetic field', *Physical Review Letters*, 100(23), 230403

Floro, J A, Chason, E, Sinclair, M B, Freund, L B and Lucadamo, G A (1998a) 'Dynamic self-organization of strained islands during SiGe epitaxial growth', *Applied Physics Letters*, 73(7), 951–953

Floro, J A, Lucadamo, G A, Chason, E, Freund, L B, Sinclair, M, Twesten, R D and Hwang, R Q (1998b) 'SiGe island shape transitions induced by elastic repulsion', *Physical Review Letters*, 80(21), 4717–4720

Friedland, K J, Siddiki, A, Hey, R, Kostial, H, Riedel, A and Maude, D K (2009) 'Quantum Hall effect in a high-mobility two-dimensional electron gas on the surface of a cylinder', *Physical Review B*, 79(12), 125320

Friesen, M, Rugheimer, P, Savage, D E, Lagally, M G, van der Weide, D W, Joynt, R and Eriksson, M A (2003) 'Practical design and simulation of silicon-based quantum-dot qubits', *Physical Review B*, 67(12), 121301

Fromherz, T, Stangl, J, Lechner, R T, Wintersberger, E, Bauer, G, Holý, V, Dais, C, Müller, E, Sigg, H, Solak, H H and Grützmacher, D (2009) '3D SiGe Quantum dot crystals: structural characterization and electronic coupling', *International Journal of Modern Physics B*, 23(12–13), 2836–2841

Fujikawa, Y, Akiyama, K, Nagao, T, Sakurai, T, Lagally, M G, Hashimoto, T, Morikawa, Y and Terakura, K (2002) 'Origin of the stability of Ge(105) on Si: a new structure model and surface strain relaxation', *Physical Reviews Letters*, 88(17), 176101

Gatti, R, Marzegalli, A, Zinovyev, V A, Montalenti, F and Miglio, L (2008) 'Modeling the plastic relaxation onset in realistic SiGe islands on Si(001)', *Physical Review B*, 78(18), 184104

Goldfarb, I, Hayden, P T, Owen, J H G and Briggs, G A D (1997) 'Competing growth mechanisms of Ge/Si(001) coherent clusters', *Physical Review B*, 56(16), 10459–10468

Golod, S V, Prinz, V Y, Mashanov, V I and Gutakovsky, A K (2001) 'Fabrication of conducting GeSi/Si micro- and nanotubes and helical microcoils', *Semiconductor Science and Technology*, 16(3), 181–185

Grinfeld, M (1994) '2-Dimensional islanding atop stressed solid helium and epitaxial-films', *Physical Review B*, 49(12), 8310–8319

Grundmann, M. (2003) 'Nanoscroll formation from strained layer heterostructures', *Applied Physics Letters*, 83(12), 2444–2446

Grützmacher, D, Fromherz, T, Dais, C, Stangl, J, Müller, E, Ekinci, Y, Solak, H H, Sigg, H, Lechner, R T, Wintersberger, E, Birner, S, Holý, V and Bauer, G (2007) 'Three-dimensional Si/Ge quantum dot crystals', *Nano Letters*, 7(10), 3150–3156

Hammar, M, LeGoues, F K, Tersoff, J, Reuter, M C and Tromp, R M (1996) 'In situ ultrahigh vacuum transmission electron microscopy studies of hetero-epitaxial growth 1 Si(001)/Ge', *Surface Science*, 349(2), 129–144

Harman, T C, Taylor, P J, Walsh, M P and LaForge, B E (2002) 'Quantum dot superlattice thermoelectric materials and devices', *Science*, 297(5590), 2229–2232

Herman, M A (1999) 'Silicon-based heterostructures: Strained-layer growth by molecular beam epitaxy', *Crystal Research and Technology*, 34(5–6), 583–595

Hesse, A, Stangl, J, Holý, V, Roch, T, Bauer, G, Schmidt, O G, Denker, U and Struth, B (2002) 'Effect of overgrowth on shape, composition, and strain of SiGe islands on Si(001)', *Physical Review B*, 66(8), 085321

Hicks, L D and Dresselhaus, M S (1993a) 'Effect of quantum-well structures on the thermoelectric figure of merit', *Physical Review B*, 47(19), 12727–12731

Hicks, L D and Dresselhaus, M S (1993b) 'Thermoelectric figure of merit of a one-dimensional conductor', *Physical Review B*, 47(24), 16631–16634

Hicks, L D, Harman, T C and Dresselhaus, M S (1993) 'Use of quantum-well superlattices to obtain a high figure of merit from nonconventional thermoelectric-materials', *Applied Physics Letters*, 63(23), 3230–3232

Hochbaum, A I, Chen, R K, Delgado, R D, Liang, W J, Garnett, E C, Najarian, M, Majumdar, A and Yang, P D (2008) 'Enhanced thermoelectric performance of rough silicon nanowires', *Nature*, 451(7175), 163–168

Holý, V, Springholz, G, Pinczolits, M and Bauer, G (1999) 'Strain induced vertical and lateral correlations in quantum dot superlattices', *Physical Review Letters*, 83(2), 356–359

Holý, V, Stangl, J, Fromherz, T, Lechner, R T, Wintersberger, E, Bauer, G, Dais, C, Müller, E and Grützmacher, D (2009) 'X-ray diffraction investigation of a three-dimensional Si/SiGe quantum dot crystal', *Physical Review B*, 79(3), 035324

Horn-von Hoegen, M, Müller, B H and Alfalou, A (1994) 'Strain relief by microroughness in surfactant-mediated growth of Ge on Si(001)', *Physical Review B*, 50(16), 11640–11652

Horn-von Hoegen, M, Müller, B H, Grabosch, T and Kury, P (2004) 'Strain relief during Ge hut cluster formation on Si(001) studied by high-resolution LEED and surface-stress-induced optical deflection', *Physical Review B*, 70(23), 235313

Houlet, L F, Shin, W, Nishibori, M, Izu, N, Itoh, T and Matsubara, L (2008) 'Fabrication and performance of free-standing hydrogen gas sensors', *Sensors and Actuators B – Chemical*, 129(1), 1–9

Huang, M H, Boone, C, Roberts, M, Savage, D E, Lagally, M G, Shaji, N, Qin, H, Blick, R, Nairn, J A and Liu, F (2005) 'Nanomechanical architecture of strained bilayer thin films: From design principles to experimental fabrication', *Advanced Materials*, 17(23), 2860–2864

Huh, D, Mills, K L, Zhu, X Y, Burns, M A, Thouless, M D and Takayama, S (2007) 'Tuneable elastomeric nanochannels for nanofluidic manipulation', *Nature Materials*, 6(6), 424–428

Hull, R, Graham, J F, Kubis, A J, Portavoce, A and Ross, F M (2007) 'Assembly and analysis of ordered semiconductor quantum dot arrays by focused ion beam nanofabrication and tomography', *Microscopy and Microanalysis*, 13, 176–177

Ismail, K, Meyerson, B S and Wang, P J (1991) 'High electron-mobility in modulation-doped Si/SiGe', *Applied Physics Letters*, 58(19), 2117–2119

Ito, K and Takahashi K (1968) 'Epitaxial growth of Ge layers on si substrates by vacuum evaporation', *Japanese Journal of Applied Physics*, 7(8), 821–826

Jesson, D E, Chen, G, Chen, K M and Pennycook, S J (1998) 'Self-limiting growth of strained faceted islands', *Physical Review Letters*, 80(23), 5156–5159

Jesson, D E, Kästner, M and Voigtländer, B (2000) 'Direct observation of subcritical fluctuations during the formation of strained semiconductor islands', *Physical Review Letters*, 84(2), 330–333

Jin, G, Liu, J L and Wang, K L (2003) 'Temperature effect on the formation of uniform self-assembled Ge dots', *Applied Physics Letters*, 83(14), 2847–2849

Kamins, T I and Williams, R S (1997) 'Lithographic positioning of self-assembled Ge islands on Si(001)', *Applied Physics Letters*, 71(9), 1201–1203

Kammler, M, Hull, R, Reuter, M C and Ross, F M (2003) 'Lateral control of self-assembled island nucleation by focused-ion-beam micropatterning', *Applied Physics Letters*, 82(7), 1093–1095

Kane, B E (1998) 'A silicon-based nuclear spin quantum computer', *Nature*, 393(6681), 133–137

Kar, G S, Kiravittaya, S, Stoffel, M and Schmidt, O G (2004) 'Material distribution across the interface of random and ordered island arrays', *Physical Review Letters*, 93(24), 246103

Kasper, E, Herzog, H J and Kibbel, H (1975) 'One-dimensional SiGe superlattice grown by UHV epitaxy', *Applied Physics*, 8(3), 199–205

Kästner, M and Voigtländer, B (1999) 'Kinetically self-limiting growth of Ge islands on Si(001)', *Physical Review Letters*, 82(13), 2745–2748

Katsaros, G, Tersoff, J, Stoffel, M, Rastelli, A, Acosta-Diaz, P, Kar, G S, Costantini, G, Schmidt, O G and Kern, K (2008) 'Positioning of strained islands by interaction with surface nanogrooves', *Physical Review Letters*, 101(9), 096103

Katsaros, G, Spathis, P, Stoffel, M, Fournel, F, Mongillo, M, Bouchiat, V, Lefloch, F, Rastelli, A, Schmidt, O G and De Franceschi, S (2010) 'Hybrid superconductor–semiconductor devices made from self-assembled SiGe nanocrystals on silicon', *Nature Nanotechnology*, 5, 458–464

Kienzle, O, Ernst, F, Ruhle, M, Schmidt, O G and Eberl, K (1999) 'Germanium 'quantum dots' embedded in silicon: Quantitative study of self-alignment and coarsening', *Applied Physics Letters*, 74(2), 269–271

Kim, E S, Usami, N and Shiraki, Y (1998) 'Control of Ge dots in dimension and position by selective epitaxial growth and their optical properties', *Applied Physics Letters*, 72(13), 1617–1619

Kim, H J, Zhao, Z M and Xie, Y H (2003) 'Three-stage nucleation and growth of Ge self-assembled quantum dots grown on partially relaxed SiGe buffer layers', *Physical Review B*, 68(20), 205312

Kim, W, Zide, J, Gossard, A, Klenov, D, Stemmer, S, Shakouri, A and Majumdar, A (2006) 'Thermal conductivity reduction and thermoelectric figure of merit increase by embedding nanoparticles in crystalline semiconductors', *Physical Review Letters*, 96(4), 045901

Kiravittaya, S, Rastelli, A and Schmidt, O G (2005) 'Self-assembled InAs quantum dots on patterned GaAs(001) substrates: Formation and shape evolution', *Applied Physics Letters*, 87(24), 243112

Koh, Y K, Vineis, C J, Calawa, S D, Walsh, M P and Cahill, D G (2009) 'Lattice thermal conductivity of nanostructured thermoelectric materials based on PbTe', *Applied Physics Letters*, 94, 153101

Köhler, U, Jusko, O, Müller, B, Horn-von Hoegen, M and Pook, M (1992) 'Layer-by-layer growth of germanium on Si(100) – Strain-induced morphology and the influence of surfactants', *Ultramicroscopy*, 42, 832–837

Kohmoto, S, Nakamura, H, Ishikawa, T and Asakawa, K (1999) 'Site-controlled self-organization of individual InAs quantum dots by scanning tunneling probe-assisted nanolithography', *Applied Physics Letters*, 75(22), 3488–3490

Lang, C, Kodambaka, S, Ross, F M and Cockayne, D J H (2006) 'Real time observation of GeSi/Si(001) island shrinkage due to surface alloying during Si capping', *Physical Review Letters*, 97(22), 226104

Ledentsov, N N, Shchukin, V A, Grundmann, M, Kirstaedter, N, Bohrer, J, Schmidt, O, Bimberg, D, Ustinov, V M, Egorov, A Y, Zhukov, A E, Kopev, P S, Zaitsev, S V, Gordeev, N Y, Alferov, Z I, Borovkov, A I, Kosogov, A O, Ruvimov, S S, Werner, P, Gosele, U and Heydenreich, J (1996) 'Direct formation of vertically coupled quantum dots in Stranski–Krastanow growth', *Physical Review B*, 54(12), 8743–8750

Lee, M L and Venkatasubramanian, R (2008) 'Effect of nanodot areal density and period on thermal conductivity in SiGe/Si nanodot superlattices', *Applied Physics Letters*, 92(5), 053112

LeGoues, F K, Reuter, M C, Tersoff, J, Hammar, M and Tromp, R M (1994) 'Cyclic growth of strain-relaxed islands', *Physical Review Letters*, 73(2), 300–303

LeGoues, F K, Hammar, M, Reuter, M C and Tromp, R M (1996) 'In situ TEM study of the growth of Ge on Si(111)', *Surface Science*, 349(3), 249–266

Li, X L (2008) 'Strain induced semiconductor nanotubes: from formation process to device applications', *Journal of Physics D – Applied Physics*, 41(19), 193001

Lin, J H, Wu, Y Q, Cui, J, Fan, Y L, Yang, X J, Jiang, Z M, Chen, Y and Zou, J (2009) 'Formation of planar defects over GeSi islands in Si capping layer grown at low temperature', *Journal of Applied Physics*, 105(2), 024307

Liu, J L, Khitun, A, Wang, K L, Borca-Tasciuc, T, Liu, W L, Chen, G and Yu, D P (2001) 'Growth of Ge quantum dot superlattices for thermoelectric applications', *Journal of Crystal Growth*, 227, 1111–1115

Liu, J L, Khitun, A, Wang, K L, Liu, W L, Chen, G, Xie, Q H and Thomas, S G (2003) 'Cross-plane thermal conductivity of self-assembled Ge quantum dot superlattices', *Physical Review B*, 67(16), 165333

Liu, X H, Ross, F M and Schwarz, K W (2000) 'Dislocated epitaxial islands', *Physical Review Letters*, 85(19), 4088–4091

Malachias, A, Mei, Y F, Annabattula, R K, Deneke, C, Onck, P R and Schmidt, O G (2008) 'Wrinkled-up nanochannel networks: Long-range ordering, scalability, and X-ray investigation', *ACS Nano*, 2(8), 1715–1721

Maradudin, A A and Wallis, R F (1980) 'Elastic interactions of point-defects in a semi-infinite medium', *Surface Science*, 91(2–3), 423–439

Marchetti, R, Montalenti, F, Miglio, L, Capellini, G, De Seta, M and Evangelisti, F (2005) 'Strain-induced ordering of small Ge islands in clusters at the surface of multilayered Si–Ge nanostructures', *Applied Physics Letters*, 87(26), 261919

Marzegalli, A, Zinovyev, V A, Montalenti, F, Rastelli, A, Stoffel, M, Merdzhanova, T, Schmidt, O G and Miglio, L (2007) 'Critical shape and size for dislocation nucleation in $Si_{1-x}Ge_{x}$ islands on Si(001)', *Physical Review Letters*, 99(23), 235505

Mateeva, E, Sutter, P, Bean, J C and Lagally, M G (1997) 'Mechanism of organization of three-dimensional islands in SiGe/Si multilayers', *Applied Physics Letters*, 71(22), 3233–3235

Matei, D, Sanduijav, B, Chen, G, Hesser, G and Springholz, G (2009) 'Molecular beam epitaxy of Si/Ge nanoislands on stripe-patterned Si (001) substrates with different stripe orientations', *Journal of Crystal Growth*, 311(7), 2220–2223

McKay, M R, Venables, J A and Drucker, J (2008) 'Kinetically suppressed Ostwald ripening of Ge/Si(100) hut clusters', *Physical Review Letters*, 101(21), 216104

Medeiros-Ribeiro, G, Kamins, T I, Ohlberg, D A A and Williams, R S (1998a) 'Annealing of Ge nanocrystals on Si(001) at 550°C: Metastability of huts and the stability of pyramids and domes', *Physical Review B*, 58(7), 3533–3536

Medeiros-Ribeiro, G, Bratkovski, A M, Kamins, T I, Ohlberg, D A A and Williams, R S (1998b) 'Shape transition of germanium nanocrystals on a silicon (001) surface from pyramids to domes', *Science*, 279(5349), 353–355

Mei, Y F, Thurmer, D J, Cavallo, F, Kiravittaya, S and Schmidt, O G (2007a) 'Semiconductor sub-micro-/nanochannel networks by deterministic layer wrinkling', *Advanced Materials*, 19(16), 2124–2128

Mei, Y F, Kiravittaya, S, Benyoucef, M, Thurmer, D J, Zander, T, Deneke, C, Cavallo, F, Rastelli, A and Schmidt, O G (2007b) 'Optical properties of a wrinkled nanomembrane with embedded quantum well', *Nano Letters*, 7(6), 1676–1679

Mei, Y F, Kiravittaya, S, Harazim, S and Schmidt, O G (2010) 'Principles and applications of micro- and nanoscale wrinkles', *Materials Science and Engineering: R: Reports* 70(3-6), 209–224

Merdzhanova, T, Kiravittaya, S, Rastelli, A, Stoffel, M, Denker, U and Schmidt, O G (2006) 'Dendrochronology of strain-relaxed islands', *Physical Review Letters*, 96(22), 226103

Meyer, G J, Dias, N L, Blick, R H and Knezevic, I (2007) 'Magnetotransport in nonplanar SiGe/Si nanomembranes', *IEEE Transactions on Nanotechnology*, 6(4), 446–450

Michler, P, Kiraz, A, Becher, C, Schoenfeld, W V, Petroff, P M, Zhang, L D, Hu, E and Imamoglu, A (2000) 'A quantum dot single-photon turnstile device', *Science*, 290, 2282–2285

Mizuno, T, Sugiyama, N, Tezuka, T and Takagi, S (2002) 'Relaxed SiGe-on-insulator substrates without thick SiGe buffer layers', *Applied Physics Letters*, 80(4), 601–603

Mo, Y W, Savage, D E, Swartzentruber, B S and Lagally, M G (1990) 'Kinetic pathway in Stranski–Krastanov growth of Ge on Si(001)', *Physical Review Letters*, 65(8), 1020–1023

Montalenti, F, Marzegalli, A, Capellini, G, de Seta, M and Miglio, L (2007) 'Vertical and lateral ordering of Ge islands grown on Si(001): theory and experiments', *Journal of Physics – Condensed Matter*, 19(22), 225001

Motta, N, Sgarlata, A, Rosei, F, Szkutnik, P D, Nufris, S, Scarselli, M and Balzarotti, A (2003) 'Controlling the quantum dot nucleation site', *Materials Science and Engineering B – Solid State Materials for Advanced Technology*, 101(1–3), 77–88

Myronov, M, Sawano, K, Shiraki, Y, Mouri, T and Itoh, K M (2007) 'Observation of two-dimensional hole gas with mobility and carrier density exceeding those of two-dimensional electron gas at room temperature in the SiGe heterostructures', *Applied*

Physics Letters, 91(8), 082108

Nakata, Y, Sugiyama, Y, Futatsugi, T and Yokoyama, N (1997) 'Self-assembled structures of closely stacked InAs islands grown on GaAs by molecular beam epitaxy', *Journal of Crystal Growth*, 175, 713–719

Narayanamurti, V (1987) 'Artificially structured thin-film materials and interfaces', *Science*, 235(4792), 1023–1028

Ni, Y, Soh, A K and He, L H (2004) 'Strain-mediated patterning of surface nanostructure by the subsurface island array', *Journal of Crystal Growth*, 269(2–4), 262–269

Nikishkov, G P (2003) 'Curvature estimation for multilayer hinged structures with initial strains', *Journal of Applied Physics*, 94(8), 5333–5336.

Nötzel, R, Temmyo, J and Tamamura, T (1994) 'Self-organized growth of strained InGaAs Quantum disks', *Nature*, 369(6476), 131–133

Ocampo, J M Z, Vaccaro, P O, Fleischmann, T, Wang, T S, Kubota, K, Aida, T, Ohnishi, T, Sugimura, A, Izumoto, R, Hosoda, M and Nashima, S (2003) 'Optical actuation of micromirrors fabricated by the micro-origami technique', *Applied Physics Letters*, 83(18), 3647–3649

Paul, D J (2004) 'Si/SiGe heterostructures: from material and physics to devices and circuits', *Semiconductor Science and Technology*, 19(10), R75–R108

People, R and Bean, J C (1986) 'Band alignments of coherently strained $\mathrm{Ge}_x\mathrm{si}_{1-x}$/Si heterostructures on ⟨001⟩ $\mathrm{Ge}_y\mathrm{Si}_{1-y}$ substrates', *Applied Physics Letters*, 48(8), 538–540

Pernot, G, Stoffel, M, Savic, I, Pezzoli, F, Chen, P, Savelli, G, Jacquot, A, Schumann, J, Denker, U, Mönch, I, Deneke, C, Schmidt, O G, Rampnoux, J M, Wang, S, Plissonnier, M, Rastelli, A, Dilhaire, S and Mingo N (2010) 'Precise control of thermal conductivity at the nanoscale through individual phonon-scattering barriers', *Nature Materials*, 9, 491–495

Persichetti, L, Sgarlata, A, Fanfoni, M and Balzarotti, A (2010) 'Shaping Ge islands on Si(001) surfaces with misorientation angle', *Physical Review Letters*, 104(3), 036104

Pezzoli, F, Stoffel, M, Merdzhanova, T, Rastelli, A and Schmidt, O G (2009) 'Alloying and strain relaxation in SiGe islands grown on pit-patterned Si(001) substrates probed by nanotomography', *Nanoscale Research Letters*, 4(9), 1073–1077

Portavoce, A, Hull, R, Reuter, M C and Ross, F M (2007) 'Nanometer-scale control of single quantum dot nucleation through focused ion-beam implantation', *Physical Review B*, 76(23), 235301

Prinz, V Y, Seleznev, V A, Gutakovsky, A K, Chehovskiy, A V, Preobrazhenskii, V V, Putyato, M A and Gavrilova, T A (2000) 'Free-standing and overgrown InGaAs/GaAs nanotubes, nanohelices and their arrays', *Physica E*, 6(1–4), 828–831

Prinz, V Y, Grützmacher, D, Beyer, A, David, C, Ketterer, B and Deckardt, E (2001) 'A new technique for fabricating three-dimensional micro- and nanostructures of various shapes', *Nanotechnology*, 12(4), 399–402

Raiteri, P, Migas, D B, Miglio, L, Rastelli, A and von Känel, H (2002) 'Critical role of the surface reconstruction in the thermodynamic stability of {105} Ge pyramids on Si(001)', *Physical Review Letters*, 88(25), 256103

Rastelli, A and von Känel, H (2002) 'Surface evolution of faceted islands', *Surface Science*, 515(2–3), L493–L498

Rastelli, A, Kummer, M and von Känel, H (2001) 'Reversible shape evolution of Ge islands on Si(001)', *Physical Review Letters*, 87(25), 256101

Rastelli, A, Müller, E and von Känel, H (2002) 'Shape preservation of Ge/Si(001) islands during Si capping', *Applied Physics Letters*, 80(8), 1438–1440

Rastelli, A, von Känel, H, Spencer, B J and Tersoff, J (2003) 'Prepyramid-to-pyramid transition of SiGe islands on Si(001)', *Physical Review B*, 68(11), 115301

Rastelli, A, Stoffel, M, Tersoff, J, Kar, G S and Schmidt, O G (2005) 'Kinetic evolution and equilibrium morphology of strained islands', *Physical Review Letters*, 95(2), 026103

Rastelli, A, Stoffel, M, Denker, U, Merdzhanova, T and Schmidt, O G (2006) 'Strained SiGe islands on Si(001): Evolution, motion, dissolution, and plastic relaxation', *Physica Status Solidi A – Applications and Materials Science*, 203(14), 3506–3511

Richard, M I, Schülli, T U, Renaud, G, Wintersberger, E, Chen, G, Bauer, G and Holý, V (2009) 'In situ X-ray scattering study on the evolution of Ge island morphology and relaxation for low growth rate: Advanced transition to superdomes', *Physical Review B*, 80(4), 045313

Ross, F M, Tromp, R M and Reuter, M C (1999) 'Transition states between pyramids and domes during Ge/Si island growth', *Science*, 286(5446), 1931–1934

Salling, C T and Lagally, M G (1994) 'Fabrication of atomic-scale structures on Si(001) surfaces', *Science*, 265(5171), 502–506

Sanduijav, B, Matei, D, Chen, G and Springholz, G (2009) 'Shape transitions and island nucleation for Si/Ge molecular beam epitaxy on stripe-patterned Si (001) substrate', *Physical Review B*, 80(12), 125329

Schäffler, F (1997) 'High-mobility Si and Ge structures', *Semiconductor Science and Technology*, 12(12), 1515–1549

Schell-Sorokin, A J and Tromp, R M (1990) 'Mechanical stresses in (sub)monolayer epitaxial-films', *Physical Review Letters*, 64(9), 1039–1042

Schittenhelm, P, Abstreiter, G, Darhuber, A, Bauer, G, Werner, P and Kosogov, A (1997) 'Growth of self-assembled homogeneous SiGe-dots on Si(100)', *Thin Solid Films*, 294(1–2), 291–295

Schmidt O G (2007), *Lateral Alignment of Epitaxial Quantum Dots*, Berlin, Springer

Schmidt, O G and Eberl, K (2001a) 'Nanotechnology – Thin solid films roll up into nanotubes', *Nature*, 410(6825), 168

Schmidt, O G and Eberl, K (2001b) 'Self-assembled Ge/Si dots for faster field-effect transistors', *IEEE Transactions on Electron Devices*, 48(6), 1175–1179

Schmidt, O G and Jin-Phillipp, N Y (2001) 'Free-standing SiGe-based nanopipelines on Si(001) substrates', *Applied Physics Letters*, 78(21), 3310–3312

Schmidt, O G, Kienzle, O, Hao, Y, Eberl, K and Ernst, F (1999) 'Modified Stranski–Krastanov growth in stacked layers of self-assembled islands', *Applied Physics Letters*, 74(9), 1272–1274

Schmidt, O G, Denker, U, Eberl, K, Kienzle, O and Ernst, F (2000a) 'Effect of overgrowth temperature on the photoluminescence of Ge/Si islands', *Applied Physics Letters*, 77(16), 2509–2511

Schmidt, O G, Jin-Phillipp, N Y, Lange, C, Denker, U, Eberl, K, Schreiner, R, Grabeldinger, H and Schweizer, H (2000b) 'Long-range ordered lines of self-assembled Ge islands on a flat Si(001) surface', *Applied Physics Letters*, 77(25), 4139–4141

Schmidt, O G, Eberl, K and Rau, Y (2000c) 'Strain and band-edge alignment in single and multiple layers of self-assembled Ge/Si and GeSi/Si islands', *Physical Review B*, 62(24), 16715–16720

Schmidt, O G, Schmarje, N, Deneke, C, Müller, C and Jin-Phillipp, N Y (2001) 'Three-dimensional nano-objects evolving from a two-dimensional layer technology', *Advanced Materials*, 13(10), 756–759

Schmidt, O G, Deneke, C, Manz, Y M and Müller, C (2002a) 'Semiconductor tubes, rods and rings of nanometer and micrometer dimension', *Physica E – Low-Dimensional Systems and Nanostructures*, 13(2–4), 969–973

Schmidt, O G, Denker, U, Dashiell, M, Jin-Phillipp, N Y, Eberl, K, Schreiner, R, Grabeldinger, H, Schweizer, H, Christiansen, S and Ernst, F (2002b) 'Laterally aligned Ge/Si islands: a new concept for faster field-effect transistors', *Materials Science and Engineering B – Solid State Materials for Advanced Technology*, 89(1–3), 101–105

Schmidt, O G, Kiravittaya, S, Nakamura, Y, Heidemeyer, H, Songmuang, R, Müller, C, Jin-Phillipp, N Y, Eberl, K, Wawra, H, Christiansen, S, Grabeldinger, H and Schweizer, H (2002c) 'Self-assembled semiconductor nanostructures: Climbing up the ladder of order', *Surface Science*, 514(1–3), 10–18

Schülli, T U, Stoffel, M, Hesse, A, Stangl, J, Lechner, R T, Wintersberger, E, Sztucki, M, Metzger, T H, Schmidt, O G and Bauer, G (2005) 'Influence of growth temperature on interdiffusion in uncapped SiGe-islands on Si(001) determined by anomalous X-ray diffraction and reciprocal space mapping', *Physical Review B*, 71(3), 035326

Schülli, T U, Vastola, G, Richard, M I, Malachias, A, Renaud, G, Uhlik, F, Montalenti, F, Chen, G, Miglio, L, Schäffler, F and Bauer, G (2009) 'Enhanced relaxation and intermixing in Ge islands grown on pit-patterned Si(001) substrates', *Physical Review Letters*, 102(2), 025502

Sgarlata, A, Szkutnik, P D, Balzarotti, A, Motta, N and Rosei, F (2003) 'Self-ordering of Ge islands on step-bunched Si(111) surfaces', *Applied Physics Letters*, 83(19), 4002–4004

Shaji, N, Qin, H, Blick, R H, Klein, L J, Deneke, C and Schmidt, O G (2007) 'Magnetotransport through two-dimensional electron gas in a tubular geometry', *Applied Physics Letters*, 90(4), 042101

Shakouri, A (2006) 'Nanoscale thermal transport and microrefrigerators on a chip', *Proceedings of the IEEE*, 94(8), 1613–1638

Shaleev, M V, Novikov, A V, Kuznetsov, O A, Yablonsky, A N, Vostokov, N V, Drozdov, Y N, Lobanov, D N and Krasilnik, Z F (2005) 'Ge self-assembled islands grown on SiGe/Si(001) relaxed buffer layers', *Materials Science and Engineering B – Solid State Materials for Advanced Technology*, 124, 466–469

Shchukin, V A and Bimberg, D (1999) 'Spontaneous ordering of nanostructures on crystal surfaces', *Reviews of Modern Physics*, 71(4), 1125–1171

Shchukin, V A, Ledentsov, N N, Kopev, P S and Bimberg, D (1995) 'Spontaneous ordering of arrays of coherent strained islands', *Physical Review Letters*, 75(16), 2968–2971

Shields, A J (2007) 'Semiconductor quantum light sources', *Nature Photonics*, 1(4), 215–223

Shiryaev, S Y, Jensen, F, Hansen, J L, Petersen, J W and Larsen, A N (1997) 'Nanoscale structuring by misfit dislocations in $Si_{1-x}Ge_x$/Si: Epitaxial systems', *Physical Review Letters*, 78(3), 503–506

Solomon, G S, Trezza, J A, Marshall, A F and Harris, J S (1996) 'Vertically aligned and electronically coupled growth induced InAs islands in GaAs', *Physical Review Letters*, 76(6), 952–955

Songmuang, R, Deneke, C and Schmidt, O G (2006) 'Rolled-up micro- and nanotubes from single-material thin films', *Applied Physics Letters*, 89(22), 223109

Songmuang, R, Rastelli, A, Mendach, S, Deneke, C and Schmidt, O G (2007a) 'From rolled-up Si microtubes to SiO_x/Si optical ring resonators', *Microelectronic Engineering*, 84(5–8), 1427–1430

Songmuang, R, Rastelli, A, Mendach, S and Schmidt, O G (2007b) 'SiO_x/Si radial superlattices and microtube optical ring resonators', *Applied Physics Letters*, 90(9), 091905

Srolovitz, D J (1989) 'On the stability of surfaces of stressed solids', *Acta Metallurgica*, 37(2), 621–625

Stangl, J, Holý, V and Bauer, G (2004) 'Structural properties of self-organized semiconductor nanostructures', *Reviews of Modern Physics*, 76(3), 725–783

Stoffel, M, Rastelli, A, Tersoff, J, Merdzhanova, T and Schmidt, O G (2006) 'Local equilibrium and global relaxation of strained SiGe/Si(001) layers', *Physical Review B*, 74(15), 155326

Stringfellow, G B (1982) 'Epitaxy', *Reports on Progress in Physics*, 45(5), 469–525

Sugii, N, Yamaguchi, S and Washio, K (2002) 'SiGe-on-insulator substrate fabricated by melt solidification for a strained-silicon complementary metal-oxide-semiconductor', *Journal of Vacuum Science and Technology B*, 20(5), 1891–1896

Sutter, E, Sutter, P and Bernard, J E (2004) 'Extended shape evolution of low mismatch $Si_{1-x}Ge_x$ alloy islands on Si(100)', *Applied Physics Letters*, 84(13), 2262–2264

Sutter, P and Lagally, M G (2000) 'Nucleationless three-dimensional island formation in low-misfit heteroepitaxy', *Physical Review Letters*, 84(20), 4637–4640

Teichert, C (2002) 'Self-organization of nanostructures in semiconductor heteroepitaxy', *Physics Reports – Review Section of Physics Letters*, 365(5–6), 335–432

Teichert, C, Lagally, M G, Peticolas, L J, Bean, J C and Tersoff, J (1996) 'Stress-induced self-organization of nanoscale structures in SiGe/Si multilayer films', *Physical Review B*, 53(24), 16334–16337

Teichert, C, Bean, J C and Lagally, M G (1998) 'Self-organized nanostructures in $Si_{1-x}Ge_x$ films on Si(001)', *Applied Physics A – Materials Science and Processing*, 67(6), 675–685

Tersoff, J and LeGoues, F K (1994) 'Competing relaxation mechanisms in strained layers', *Physical Review Letters*, 72(22), 3570–3573

Tersoff, J, Teichert, C and Lagally, M G (1996) 'Self-organization in growth of quantum dot superlattices', *Physical Review Letters*, 76(10), 1675–1678

Tersoff, J, Spencer, B J, Rastelli, A and von Känel, H (2002) 'Barrierless formation and faceting of SiGe islands on Si(001)', *Physical Review Letters*, 89(19), 196104

Tezuka, T, Sugiyama, N and Takagi, S (2001) 'Fabrication of strained Si on an ultrathin SiGe-on-insulator virtual substrate with a high-Ge fraction', *Applied Physics Letters*, 79(12), 1798–1800

Tomitori, M, Watanabe, K, Kobayashi, M and Nishikawa, O (1994) 'STM study of the Ge growth mode on Si(001) substrates', *Applied Surface Science*, 76(1–4), 322–328

Tromp, R M, Ross, F M and Reuter, M C (2000) 'Instability-driven SiGe island growth', *Physical Review Letters*, 84(20), 4641–4644

Tsai, N C and Sue, C Y (2007) 'Review of MEMS-based drug delivery and dosing systems', *Sensors and Actuators A – Physical*, 134(2), 555–564

Tsybeskov, L and Lockwood, D J (2009) 'Silicon–germanium nanostructures for light emitters and on-chip optical interconnects', *Proceedings of the IEEE*, 97(7), 1284–1303

Tsybeskov, L, Hirschman, K D, Duttagupta, S P, Zacharias, M, Fauchet, P M, McCaffrey, J P and Lockwood, D J (1998) 'Nanocrystalline-silicon superlattice produced by controlled recrystallization', *Applied Physics Letters*, 72(1), 43–45

Usami, N, Araki, Y, Ito, Y, Miura, M and Shiraki, Y (2000) 'Modification of the growth mode of Ge on Si by buried Ge islands', *Applied Physics Letters*, 76(25), 3723–3725

Vaccaro, P O, Kubota, K and Aida, T (2001) 'Strain-driven self-positioning of micromachined structures', *Applied Physics Letters*, 78(19), 2852–2854

Vaccaro, P O, Kubota, K, Fleischmann, T, Saravanan, S and Aida, T (2003) 'Valley-fold and mountain-fold in the micro-origami technique', *Microelectronics Journal*, 34(5–8), 447–449

Vailionis, A, Cho, B, Glass, G, Desjardins, P, Cahill, D G and Greene, J E (2000) 'Pathway for the strain-driven two-dimensional to three-dimensional transition during growth of Ge on Si(001)', *Physical Review Letters*, 85(17), 3672–3675

Vastola, G, Montalenti, F and Miglio, L (2008) 'Understanding the elastic relaxation mechanisms of strain in Ge islands on pit-patterned Si(001) substrates', *Journal of Physics – Condensed Matter*, 20(45), 454217

Venkatasubramanian, R, Siivola, E, Colpitts, T and O'Quinn, B (2001) 'Thin-film thermoelectric devices with high room-temperature figures of merit', *Nature*, 413(6856), 597–602

Vescan, L, Grimm, K, Goryll, M and Hollander, B (2000) 'Ordered nucleation of Ge islands along high index planes on Si', *Materials Science and Engineering B – Solid State Materials for Advanced Technology*, 69, 324–328

Vineis, C J, Harman, T C, Calawa, S D, Walsh, M P, Reeder, R E, Singh, R and Shakouri, A (2008) 'Carrier concentration and temperature dependence of the electronic transport properties of epitaxial PbTe and PbTe/PbSe nanodot superlattices', *Physical Review B*, 77, 235202

Voigtländer, B (2001) 'Fundamental processes in Si/Si and Ge/Si epitaxy studied by scanning tunneling microscopy during growth', *Surface Science Reports*, 43(5–8), 127–254

Vorob'ev, A B and Prinz, V Y (2002) 'Directional rolling of strained heterofilms', *Semiconductor Science and Technology*, 17(6), 614–616

Wacker, A (2002) 'Semiconductor superlattices: a model system for nonlinear transport', *Physics Reports – Review Section of Physics Letters*, 357(1), 1–111

Wang, K L, Tong, S and Kim, H J (2005) 'Properties and applications of SiGe nanodots', *Materials Science in Semiconductor Processing*, 8(1–3), 389–399

Wedler, G, Walz, J, Hesjedal, T, Chilla, E and Koch, R (1998) 'Stress and relief of misfit strain of Ge/Si(001)', *Physical Review Letters*, 80(11), 2382–2385

Wood, D M and Zunger, A (1989) 'Epitaxial effects on coherent phase-diagrams of alloys', *Physical Review B*, 40(6), 4062–4089

Xie, Q H, Madhukar, A, Chen, P and Kobayashi, N P (1995) 'Vertically self-organized InAs quantum box islands on GaAs(100)', *Physical Review Letters*, 75(13), 2542–2545

Yakimov, A I, Dvurechenskii, A V, Minkov, G M, Sherstobitov, A A, Nikiforov, A I, Bloshkin, A A, Stepina, N P, Leitao, J P, Sobolev, N A, Pereira, L and do Carmo, M C (2006) 'Hopping magnetoresistance in two-dimensional arrays of Ge/Si quantum dots', in Brom, H B, ed, *Physica Status Solidi C – Current Topics in Solid State Physics*, 3(2), 296–299

Yang, B, Liu, W L, Liu, J L, Wang, K L and Chen, G (2002a) 'Measurements of anisotropic thermoelectric properties in superlattices', *Applied Physics Letters*, 81(19), 3588–3590

Yang, B, Liu, J L, Wang, K L and Chen, G (2002b) 'Simultaneous measurements of Seebeck coefficient and thermal conductivity across superlattice', *Applied Physics Letters*, 80(10), 1758–1760

Yoo, P J and Lee, H H (2003) 'Evolution of a stress-driven pattern in thin bilayer films: Spinodal wrinkling', *Physical Review Letters*, 91(15), 154502

Zander, T, Deneke, C, Malachias, A, Mickel, C, Metzger, T H and Schmidt, O G (2009) 'Planar hybrid superlattices by compression of rolled-up nanomembranes', *Applied Physics Letters*, 053102 (3 pp)

Zhang, J J, Stoffel, M, Rastelli, A, Schmidt, O G, Jovanovic, V, Nanver, L K and Bauer, G (2007) 'SiGe growth on patterned Si(001) substrates: Surface evolution and evidence of modified island coarsening', *Applied Physics Letters*, 91(17), 173115

Zhang, J J, Rastelli, A, Groiss, H, Tersoff, J, Schäffler, F, Schmidt, O G and Bauer, G (2009) 'Shaping site-controlled uniform arrays of SiGe/Si(001) islands by in situ annealing', *Applied Physics Letters*, 95(18), 183102

Zhang, L, Deckhardt, E, Weber, A, Schonenberger, C and Grützmacher, D (2005) 'Controllable fabrication of SiGe/Si and SiGe/Si/Cr helical nanobelts', *Nanotechnology*, 16(6), 655–663

Zhang, L, Ruh, E, Grützmacher, D, Dong, L X, Bell, D J, Nelson, B J and Schonenberger, C (2006) 'Anomalous coiling of SiGe/Si and SiGe/Si/Cr helical nanobelts', *Nano Letters*, 6(7), 1311–1317

Zhang, L, Dong, L X and Nelson, B J (2008a) 'Bending and buckling of rolled-up SiGe/Si microtubes using nanorobotic manipulation', *Applied Physics Letters*, 92(24), 243102

Zhang, L, Dong, L X and Nelson, B J (2008b) 'Ring closure of rolled-up Si/Cr nanoribbons', *Applied Physics Letters*, 92(14), 143110

Zhang, L, Abbott, J J, Dong, L X, Kratochvil, B E, Bell, D and Nelson, B J (2009) 'Artificial bacterial flagella: Fabrication and magnetic control', *Applied Physics Letters*, 94(6), 064107

Zhang, P P, Nordberg, E P, Park, B N, Celler, G K, Knezevic, I, Evans, P G, Eriksson, M A and Lagally, M G (2006a) 'Electrical conductivity in silicon nanomembranes', *New Journal of Physics*, 8(9), 200

Zhang, P P, Tevaarwerk, E, Park, B N, Savage, D E, Celler, G K, Knezevic, I, Evans, P G, Eriksson, M A and Lagally, M G (2006b) 'Electronic transport in nanometre-scale silicon-on-insulator membranes', *Nature*, 439(7077), 703–706

Zhang, Y, Zeng, G H and Shakouri, A (2006) 'Silicon microrefrigerator', *IEEE Transactions on Components and Packaging Technologies*, 29(3), 570–576

Zhong, Z Y and Bauer, G (2004) 'Site-controlled and size-homogeneous Ge islands on prepatterned Si(001) substrates', *Applied Physics Letters*, 84(11), 1922–1924

Zhong, Z Y, Halilovic, A, Fromherz, T, Schäffler, F and Bauer, G (2003) 'Two-dimensional periodic positioning of self-assembled Ge islands on prepatterned Si(001) substrates', *Applied Physics Letters*, 82(26), 4779–4781

Zhong, Z, Chen, P, Jiang, Z and Bauer, G (2008) 'Temperature dependence of ordered GeSi island growth on patterned Si(001) substrates', *Applied Physics Letters*, 93(4), 043106

Zhu, J H, Brunner, K and Abstreiter, G (1998) 'Two-dimensional ordering of self-assembled Ge islands on vicinal Si(001) surfaces with regular ripples', *Applied Physics Letters*, 73(5), 620–622

Zundel, M K, Specht, P, Eberl, K, Jin Phillipp, N Y and Phillipp, F (1997) 'Structural and optical properties of vertically aligned InP quantum dots', *Applied Physics Letters*, 71(20), 2972–2974

Žutić, I, Fabian, J and Das Sarma, S (2004) 'Spintronics: Fundamentals and applications', *Reviews of Modern Physics*, 76(2), 323–410

12

Self-diffusion and dopant diffusion in germanium (Ge) and silicon–germanium (SiGe) alloys

M. UEMATSU, Keio University, Japan

Abstract: This chapter deals with self-diffusion and dopant diffusion in Ge and SiGe, including dopant segregation. The chapter begins with a general introduction to the diffusion mechanism. It then describes self-diffusion in Ge and SiGe and in addition Si–Ge interdiffusion, including isotope engineering, where the dependence of diffusivities on Ge composition and on strain is discussed. The chapter then describes the diffusion of common Group V donor and Group III acceptor dopants in Ge and SiGe.

Key words: self-diffusion in Ge and SiGe, dopant diffusion in Ge and SiGe, Si–Ge interdiffusion, isotope heterostructure, strain effect on diffusion, transient enhanced diffusion, dopant segregation.

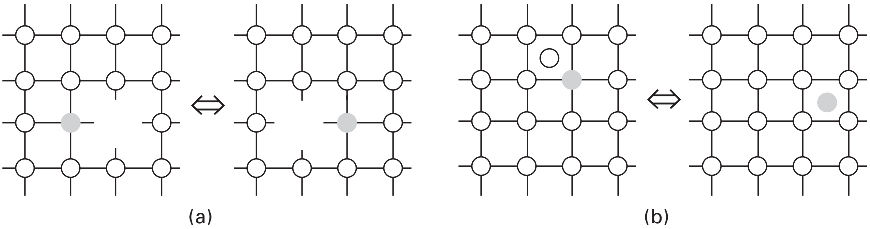
12.1 Introduction

This chapter deals with self-diffusion and dopant diffusion in Ge and SiGe, including dopant segregation. The diffusion processes are of vital importance for the formation of shallow junctions in electronic devices. Recent considerable progress in this area is highly attributed to the availability of stable isotope Si and Ge. This opens the way for direct experimental observation of self-diffusion and therefore for elucidating the mechanism of dopant diffusion that is mediated by point defects of host materials. This chapter is devoted to an up-to-date presentation of both the experimental data, including those using isotopically enriched heterostructures, and the diffusion mechanism. In Section 12.2, a general introduction to the diffusion mechanism in semiconductors is given. Section 12.3 describes self-diffusion in Ge, including isotope engineering. Next, Si and Ge diffusion in SiGe alloys (Section 12.4) and Si–Ge interdiffusion (Section 12.5) are described. The dependence of diffusivities on Ge composition and the effect of strain are discussed. Sections 12.6 and 12.7 are devoted to the diffusion of common Group V donor and Group III acceptor dopants in Ge and SiGe. In Section 12.8, dopant segregation is described. Finally, future trends are described (Section 12.9) and a short description of sources of further information and advice is given (Section 12.10).

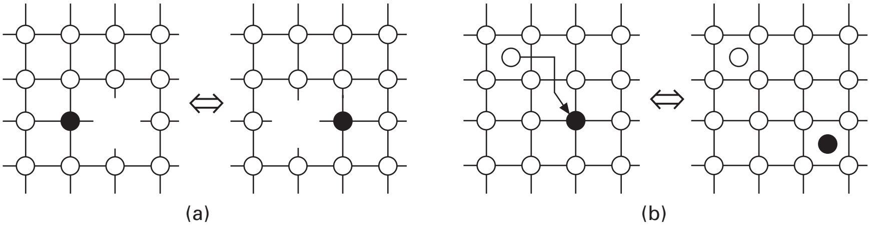
12.2 Diffusion mechanism

Self-diffusion in semiconductors is generally mediated by native point defects, namely, vacancies (V) and self-interstitials (I). Figure 12.1 illustrates (a) the vacancy and (b) the interstitial mechanisms. A diffusing atom is marked by a gray circle and it can be an isotope self-atom. In the vacancy mechanism, a diffusing atom moves by jumping into a neighboring vacancy lattice site. The diffusing atom will further move when a vacancy diffuses to occupy one of the neighboring lattice sites, which the atom jumps into. In the interstitial mechanism, a self-interstitial atom approaches an atom in a lattice site and the lattice site atom exchanges its original position with the self-interstitial. In this manner, the self-interstitial becomes a lattice site atom, whereas the original lattice site atom temporarily becomes a self-interstitial. The atom will further move when it exchanges the position with a diffusing self-interstitial.

The diffusion of substitutional dopant atoms (donors and acceptors) in semiconductors is also mediated by vacancies and self-interstitials. Figure 12.2 illustrates (a) the vacancy mechanism and (b) the interstitial mechanism, or the kick-out mechanism, where a dopant atom is marked by a black circle.



12.1 Self-diffusion mechanisms. (a) In the vacancy mechanism, the marked (isotope) atom (gray circle) moves by jumping into the neighboring vacancy. (b) In the interstitial mechanism, the marked atom exchanges its position with the self-interstitial.



12.2 Dopant diffusion mechanisms: (a) the vacancy mechanism; (b) the interstitial mechanism, or the kick-out mechanism. A dopant atom is marked by a black circle.

The vacancy-mediated dopant diffusion is basically the same as the vacancy mechanism of self-diffusion (Fig. 12.1(a)). For dopant diffusion, however, a Coulomb attractive force between a dopant atom and a neighboring vacancy may enhance the diffusion of dopant atoms compared with that of self-atoms. For this reason, vacancy-mediated dopant diffusion is normally faster than vacancy-mediated self-diffusion. The vacancy mechanism of dopant diffusion is described by the pair diffusion model:

$$(AV) \Leftrightarrow A_s + V \qquad 12.1$$

where A_s and (AV) denote substitutional dopants and the nearest-neighbor pair configuration in which A_s becomes mobile through site-exchange with vacancies, respectively.

In interstitial-mediated dopant diffusion (Fig. 12.2(b)), a self-interstitial atom *kicks-out* an immobile substitutional dopant (A_s) from its lattice site to a mobile interstitial position. The kick-out mechanism is similar to the interstitial mechanism of self-diffusion (Fig. 12.1(b)). However, interstitial dopant atoms are usually more mobile than self-interstitial atoms, and therefore, dopant diffusion via the kick-out mechanism is normally faster than self-diffusion via self-interstitials. The kick-out mechanism is described by the diffusion equation:

$$A_i \Leftrightarrow A_s + I \qquad 12.2$$

where A_i denotes interstitial dopant atoms. Interstitial-mediated dopant diffusion can also be described by dopant-interstitial pair diffusion, $(AI) \Leftrightarrow A_s + I$, which is formally equivalent to the kick-out mechanism (Eq. 12.2) in diffusion modeling. As will be described in Sections 12.3 and 12.6, vacancies are the dominant native point defects that govern self-diffusion and dopant diffusion in Ge. In contrast, both vacancies and self-interstitials simultaneously operate for self- and dopant diffusion in Si.

Another important feature of diffusion in semiconductors is that the point defects are charged and their concentration depends on the carrier concentration due to the Fermi level effect. This effect leads to the dopant concentration dependence of self-diffusion and dopant diffusion, which will be concretely described below. In the diffusion mechanisms, an atom jumps between different sites across a barrier to a new position and the jump is assisted by the thermal movement. Therefore, the diffusion coefficient, or diffusivity, is generally given as an Arrhenius expression

$$D = D_0 \exp(-H/kT) \qquad 12.3$$

In Eq. 12.3, k is the Boltzmann constant, T is the temperature, and D_0 is the pre-exponential factor. The activation enthalpy H is the sum of the enthalpies of migration and formation of the diffusion species under consideration. Overviews on diffusion fundamentals can be found in Frank *et al.* (1984)

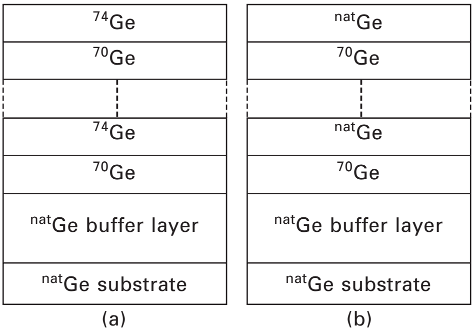
and Gösele (1988). In addition, detailed diffusion modeling and simulation are described in Uematsu (1997) and Bracht (2007).

12.3 Self-diffusion in germanium (Ge)

The most fundamental diffusion process is self-diffusion. Study of self-diffusion is a key to unveiling the properties of the native point defects and furthermore the mechanism of dopant diffusion mediated by the point defects. Isotopes are the most direct and powerful tool to investigate the self-diffusion of host materials because only the atomic mass is changing while chemical properties stay the same. The traditional technique to observe self-diffusion is to use radioactive isotope tracers with serial sectioning after diffusion. This technique, however, cannot be applied to nano-scale diffusion due to the limited lifetime of the tracers and to the instrumental broadening of the diffusion profiles associated with the sectioning processes. Recently, considerable amounts of highly enriched stable isotopes have become available and this has led to remarkable progress in self-diffusion study. Isotopically enriched superlattices (SLs) allow us to accurately determine nano-scale self-diffusivities, that is, lower self-diffusivities at lower temperatures. Experimental studies on Ge self-diffusion using Ge isotope SLs are described below.

12.3.1 Isotope engineering

Ge consists of five stable isotopes and the isotopic composition in natural Ge is ^{70}Ge (20.5%), ^{72}Ge (27.4%), ^{73}Ge (7.8%), ^{74}Ge (36.5%), and ^{76}Ge (7.8%) (Itoh *et al.*, 1993). Ge isotope SLs are composed of a stack of isotope-enriched layers, and molecular beam epitaxy (MBE) is often used for the growth. Figure 12.3 shows examples of Ge SL structures, where the thickness of each isotope layer and the number of bilayers depend on the purpose of the experiments. A $^{\mathrm{nat}}$Ge buffer layer is usually grown to achieve an atomically flat, smooth surface on top of the Ge substrate. Structure (a) consists of pure ^{70}Ge and ^{74}Ge, while (b) consists of $^{\mathrm{nat}}$Ge (natural Ge) and ^{70}Ge, where ^{74}Ge becomes a suitable diffusion marker, as will be described below. In as-grown Ge SLs, the interface between isotope layers is atomically flat and the degree of intermixing at the interface is less than two atomic monolayers (Morita *et al.*, 2000). Upon annealing, the isotopes in SLs diffuse into each other and the concentration profiles of isotopes are measured by secondary ion mass spectrometry (SIMS) to obtain the self-diffusivities (Fuchs *et al.*, 1995). Alternatively, self-diffusivities can be obtained by the observation of optical phonons in isotope SLs using Raman spectroscopy (Spitzer *et al.*, 1994; Silveira *et al.*, 1997). Lower self-diffusivities can be obtained by Raman spectroscopy compared with SIMS, in which profile broadening due to ion beam mixing cannot be avoided. In a similar manner, Si isotopes have been



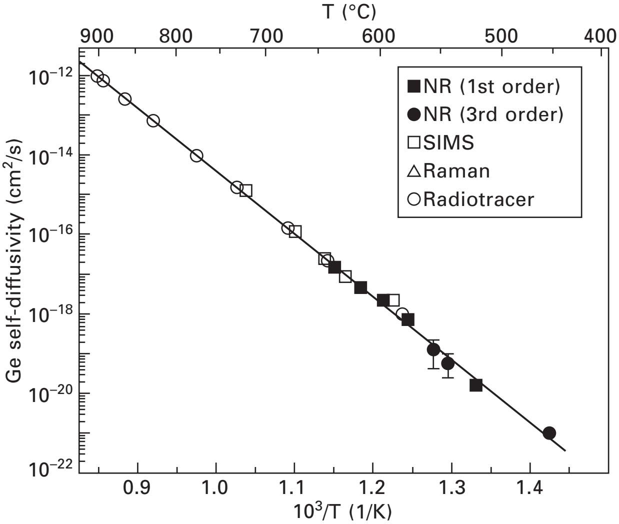
12.3 Examples of Ge SL structures: (a) ^{70}Ge/^{74}Ge, (b) ^{nat}Ge/^{70}Ge.

employed for the self-diffusion study. Si consists of three stable isotopes (^{28}Si: 92.2%, ^{29}Si: 4.7%, and ^{30}Si: 3.1%) and the isotope SL structures of ^{28}Si/^{30}Si and ^{nat}Si/^{28}Si have often been used (Shimizu *et al.*, 2007, 2009). The degree of intermixing at the Si isotope interface is also less than two atomic monolayers (Shimizu and Itoh, 2006).

Another important feature of isotope engineering is the isotopic effects on physical properties of materials. Asen-Palmer *et al.* (1997) showed that the maximum thermal conductivity measured for highly enriched ^{70}Ge (99.99%) sample was one order of magnitude higher than for natural Ge. The enhancement of thermal conductivity was also found in highly enriched ^{28}Si (Capinski *et al.*, 1997; Ruf *et al.*, 2000). These intentional manipulations of the isotopic abundances allow the development of highly thermal conductive wafers for semiconductor substrate applications. In addition, Karaiskaj *et al.* (2001) reported photoluminescence (PL) measurements of isotopically purified ^{28}Si and compared these to the same transitions in natural Si. The linewidths of the no-phonon B and P bound exciton transitions in the ^{28}Si sample were much narrower than in natural Si. The removal of the dominant broadening resulting from isotopic randomness in natural Si reveals new fine structure in the B bound exciton luminescence. Furthermore, only ^{29}Si has a non-zero nuclear spin ($I = 1/2$) among the stable Si isotopes, and only ^{73}Ge has a non-zero nuclear spin ($I = 9/2$) among the stable Ge isotopes. Fabrication of Si- and Ge-based nanostructures with controlled positioning of nuclear spins allows a wide variety of spin physics research and the realization of a quantum computer. A linear chain of ^{29}Si stable isotopes with nuclear spin $I = 1/2$ embedded in a spin-free ^{28}Si stable isotope matrix was shown to realize the 'all-silicon quantum computer' (Ladd *et al.*, 2002; Itoh, 2005).

12.3.2 Intrinsic Ge self-diffusivity

Figure 12.4 shows the self-diffusivities in intrinsic Ge obtained using isotope SLs (or heterostructures) by means of SIMS for 543–690°C annealing (Fuchs



12.4 Intrinsic Ge self-diffusivities obtained by NR (Hüger *et al.*, 2008) from the decay of the first- and third-order Bragg peaks, SIMS (Fuchs *et al.*, 1995), Raman spectroscopy (Silveira *et al.*, 1997), and radiotracer experiments (Werner *et al.*, 1985). The solid line shows the best fit based on an Arrhenius expression (adapted from Hüger *et al.*, 2008).

Table 12.1 Intrinsic self-diffusivity in Ge

H (eV)	D_0 (cm²/s)	Temperature (°C)	Technique	Reference
3.09	13.6	535–904	Radiotracer	Werner *et al.*, 1985
3.05	12.0	543–690	SIMS	Fuchs *et al.*, 1995
5.5×10^{-20} (cm²/s)		500	Raman	Silveira *et al.*, 1997
3.13	25.4	429–596	NR	Hüger *et al.*, 2008

et al., 1995) and Raman spectroscopy for 500°C (Silveira *et al.*, 1997). The diffusivities measured by radiotracer technique for higher temperatures (Werner *et al.*, 1985) are also shown in the figure. The intrinsic Ge self-diffusivity is described with one diffusion activation enthalpy. The activation enthalpies and pre-exponential factors of the self-diffusivity are summarized in Table 12.1. Taking into account the pressure dependence of self-diffusion in Ge, which revealed activation volumes for the native defect mediating self-diffusion that are typical for relaxed vacancies (Werner *et al.*, 1985), it has been concluded that the vacancy mechanism governs the self-diffusion in Ge. Quite recently, the diffusion experiment to temperatures as low as 429°C was performed by neutron reflectometry (NR) of ^{nat}Ge/^{70}Ge isotope SLs (Hüger *et al.*, 2008). The diffusivities determined by NR from the decay of the first- and third-order Bragg peaks are plotted in Fig. 12.4. The Ge self-diffusivities at the intrinsic carrier concentration (n_i) shown in the figure are accurately described by

$$D_{Ge}^{SD}\ (n_i) = 25.4 \exp(-3.13\ \text{eV}/kT)\ (\text{cm}^2/\text{s}) \qquad 12.4$$

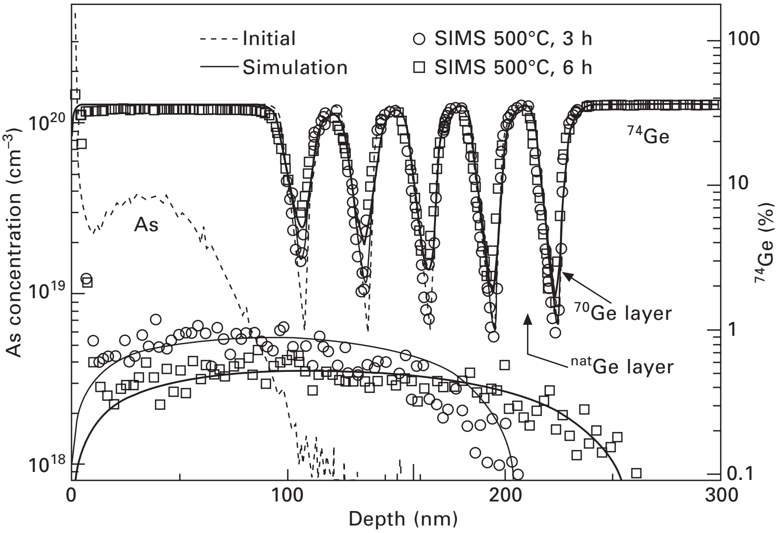
with a single activation enthalpy for temperatures between 429 and 904°C. (The diffusivity reported by Werner *et al.* (1985), which has been widely accepted and used, differs only slightly from that of Eq. 12.4.) This result confirms that only vacancies contribute self-diffusion in Ge over the whole temperature range. This is completely different from self-diffusion in Si, to which self-interstitials mainly contribute at higher temperatures ($\gtrsim$900°C) and vacancies at lower temperatures ($\lesssim$900°C) (Bracht *et al.*, 1998; Shimizu *et al.*, 2007).

12.3.3 Charge states of vacancies in Ge

The self-diffusivity varies with the carrier concentration because the concentration of point defects depends on the position of the Fermi level if they are charged (Shockley and Moll, 1960). For example, the vacancy concentration is proportional to $(n/n_i)^2$ if the vacancy is doubly negatively charged (n denotes the electron concentration). The charge states of the vacancy in Ge can be determined by the doping dependence of Ge self-diffusion. Werner *et al.* (1985) found an enhanced self-diffusion in n-type Ge and retardation in p-type Ge and therefore concluded that the vacancy acts as an acceptor. Brotzmann *et al.* (2008) observed the simultaneous diffusion of Ge and n-type dopant atoms (P, As, and Sb) in ^{nat}Ge/^{70}Ge isotope SLs. The Ge self-diffusion profiles showed an enhanced intermixing of the ^{nat}Ge/^{70}Ge layers within the dopant profiles. The doping dependence of the Ge self-diffusion was best described with the contribution of doubly negatively charged vacancies (V^{2-}). Naganawa *et al.* (2008a) also simultaneously investigated Ge self-diffusion and As diffusion in As-implanted ^{nat}Ge/^{70}Ge isotope SLs (Fig. 12.5, in which the actual interfaces between ^{nat}Ge and ^{70}Ge layers are abrupt and the smearing of the ^{74}Ge profiles is due to the SIMS artifact). Transient enhanced diffusion (TED) due to the implantation damages was not observed, as will be discussed in Section 12.6.3. A quadratic dependence of the Ge self-diffusion on the carrier concentration was observed. From a precise reproduction of the Ge and As diffusion profiles by numerical simulation, the authors showed that V^{2-} is the dominant point defect responsible for more than 95% of the self-diffusion in intrinsic Ge and that this fraction increases even further in n-type Ge. Based on these findings, the Ge self-diffusivity under extrinsic conditions is described by

$$D_{Ge}^{SD}(n) = D_{Ge}^{SD}(n_i)\ (n/n_i)^2 \qquad 12.5$$

which shows the quadratic dependence on the carrier concentration.



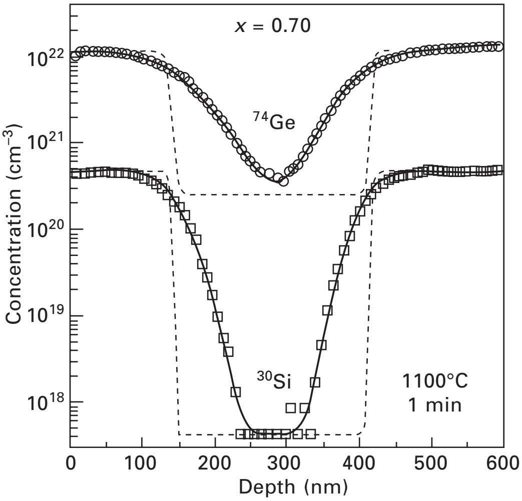
12.5 SIMS and simulated depth profiles of ^{74}Ge (upper profiles) and As (lower profiles) in the ^{nat}Ge/^{70}Ge isotope SLs implanted with As at 90 keV, 2×10^{14} cm^{-2}. A ^{nat}Ge cap layer was grown on the top of the SL. In both the upper and the lower profiles, broken lines, open circles, and open squares represent the SIMS profiles before and after annealing at 500°C for 3 and 6 h, respectively. Solid curves are the simulation results (adapted from Naganawa *et al.*, 2008a).

12.4 Self-diffusion in silicon–germanium (SiGe) alloys

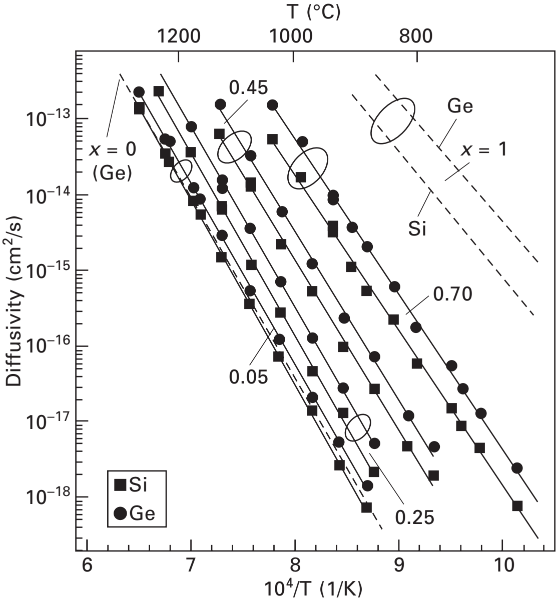
12.4.1 Ge composition dependence

Self-diffusion, or Si and Ge diffusion, in strain-relaxed SiGe alloys has been investigated by several groups. The diffusivity of Ge in polycrystalline SiGe was measured by means of a ^{71}Ge radiotracer technique (McVay and DuCharme, 1974) and those of Si and Ge in monocrystalline SiGe were measured using ^{31}Si and ^{71}Ge radiotracers (Strohm *et al.*, 2002). Zangenberg *et al.* (2001) investigated Ge diffusion in MBE-grown SiGe using ^{72}Ge isotopically enriched Ge. These results, however, do not show a consistent result on the activation enthalpy of Ge diffusion, as will be described below. Quite recently, simultaneous observation of Si and Ge diffusion using isotopically controlled $Si_{1-x}Ge_x$ was performed by Kube *et al.* (2009a, 2009b). They measured the self-diffusivities of ^{30}Si and ^{74}Ge in $^{nat}Si_{1-x}{}^{nat}Ge_x/{}^{28}Si_{1-x}{}^{70}Ge_x/{}^{nat}Si_{1-x}{}^{nat}Ge_x$ isotope structures with $x = 0.05$, 0.25, 0.45, and 0.70 at 690–1270°C. In this structure, a graded buffer layer (of thickness > 1 μm) was first grown on top of the Si substrate to reduce the threading dislocation density (in the range of 5×10^5 cm^{-2}) associated with the lattice mismatch between Si and SiGe. Figure 12.6 shows the SIMS concentration profiles of ^{30}Si and ^{74}Ge in the isotope structure with $x = 0.70$ before and after annealing at 1100°C for 1 min. For Ge diffusion in elemental Si ($x = 0$), ^{74}Ge profiles of a ^{nat}Ge buried layer (with ^{nat}Ge concentration of about 10^{20} cm^{-3}) in Si were observed. The measured diffusion profiles were fitted by a single self-diffusivity for each profile.

The temperature dependences of Si and Ge diffusivities in $Si_{1-x}Ge_x$ alloys are shown in Fig. 12.7. An Arrhenius-type temperature dependence of diffusion was observed for all Ge compositions. For comparison, the Si diffusivity in Ge ($x = 1$) (Silvestri *et al.*, 2006) and the Ge self-diffusivity (Werner *et al.*, 1985) are also shown in the figure. The activation enthalpies and pre-exponential factors of these diffusivities are summarized in Table 12.2. The results show a clear trend that the Si and Ge diffusivities in $Si_{1-x}Ge_x$ alloys increase with an increase of the Ge content, x. This trend essentially results from the increase in the number of weaker and less stiff Si–Ge and Ge–Ge bonds compared with the number of stiffer and stronger Si–Si bonds (Aubry *et al.*, 1999). In Si ($x = 0$), the Ge diffusivity almost equals the Si self-diffusivity (Bracht *et al.*, 1998). With increasing Ge content, however, the diffusion of Ge gets increasingly faster compared to that of Si and the ratio between the Ge and Si diffusivities increases. This trend is explained by an increasing contribution of vacancies to self-diffusion in $Si_{1-x}Ge_x$ with an increase of the Ge content (both vacancies and self-interstitials in Si but



12.6 SIMS concentration profiles of ^{30}Si and ^{74}Ge in the $^{nat}Si_{1-x}{}^{nat}Ge_x/^{28}Si_{1-x}{}^{70}Ge_x/^{nat}Si_{1-x}{}^{nat}Ge_x$ isotope structure with $x = 0.70$ before (broken lines) and after annealing (symbols) at 1100°C for 1 min. The solid lines are best fits based on the solution of Fick's law (adapted from Kube *et al.*, 2009a, 2009b).



12.7 Temperature dependence of Si (squares) and Ge (circles) diffusivities in $Si_{1-x}Ge_x$ alloys with $x = 0.05$, 0.25, 0.45, and 0.70. The solid lines show best fits based on an Arrhenius expression. The temperature dependence of Ge diffusivity in Si ($x = 0$) is also shown (dashed line), which almost equals Si self-diffusivity (Bracht *et al.*, 1998). For the diffusion in Ge ($x = 1$), Si diffusivity in Ge (Silvestri *et al.*, 2006) and Ge self-diffusivity (Werner *et al.*, 1985) are shown (dashed lines) (adapted from Kube *et al.*, 2009a, 2009b).

Table 12.2 Self-diffusivities of Si and Ge in $Si_{1-x}Ge_x$ alloys

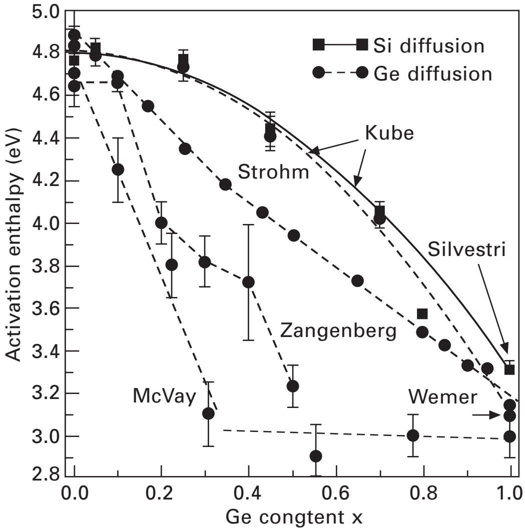
	Si		Ge		
x	H (eV)	D_0 (cm²/s)	H (eV)	D_0 (cm²/s)	Temperature (°C)
0	4.75[a]	530[a]	4.83	923	880–1270
0.05	4.82	795	4.77	915	880–1250
0.25	4.77	2233	4.71	3022	870–1225
0.45	4.44	1075	4.42	2074	800–1105
0.70	4.05	366	4.05	1126	690–1015
1	3.32[b]	42[b]	3.09[c]	13.6[c]	–

[a] Bracht *et al.*, 1998 (temperature 855–1388°C).
[b] Silvestri *et al.*, 2006 (550–900°C).
[c] Werner *et al.*, 1985 (535–904°C).
Source: Kube *et al.*, 2009a, 2009b.

almost only vacancies in Ge). The more pronounced Ge diffusion indicates that with increasing Ge content vacancies in SiGe are more likely to be attracted by Ge than by Si and therefore facilitate the Ge diffusion (Kube *et al.*, 2009a, 2009b).

12.4.2 Activation enthalpy of diffusion

The dependence of the activation enthalpies of Si and Ge diffusion on the Ge content of $Si_{1-x}Ge_x$ is illustrated in Fig. 12.8. The composition dependence of the activation enthalpy reported by Kube *et al.* (2009a, 2009b) reveals an upward bowing, while the activation enthalpies by McVay and DuCharme (1974), Zangenberg *et al.* (2001), and Strohm *et al.* (2002) show a faster decrease with the Ge content. The reason for the different trends in the composition dependence remains unsolved. According to the discussions by Kube *et al.* (2009a, 2009b), the studies of McVay and DuCharme (1974) may have suffered from the limited quality of the polycrystalline SiGe, in which grain boundary diffusion could play a significant role. The experiments performed by Zangenberg *et al.* (2001) cover only a narrow temperature range of about 100 K, and therefore the accuracy of the diffusion activation enthalpy may be limited. The experiments of Strohm *et al.* (2002) using short-lived radiotracers were also restricted to a narrow temperature range. Based on the upward bowing of the activation enthalpy, the contribution of vacancies and self-interstitials in SiGe alloys will be discussed in Section 12.7.3.



12.8 Composition dependence of the activation enthalpies of Si diffusion (squares, solid line) and Ge diffusion (circles, dashed lines) in $Si_{1-x}Ge_x$ alloys obtained by Kube *et al.* (2009a, 2009b) and the enthalpies of Si self-diffusion (square at $x = 0$) (Bracht *et al.*, 1998), Ge self-diffusion (Werner *et al.*, 1985), and Si diffusion in Ge (Silvestri *et al.*, 2006). The composition dependences of Ge diffusion in $Si_{1-x}Ge_x$ reported by McVay and DuCharme (1974), Zangenberg *et al.* (2001), and Strohm *et al.* (2002) are shown for comparison. The lines are guides for the eye (adapted from Kube *et al.*, 2009a, 2009b).

12.5 Silicon-germanium (Si–Ge) interdiffusion

Interdiffusion of Si and Ge at Si/SiGe interfaces is an important issue for a recent nano-scale metal-oxide-semiconductor field effect transistor (MOSFET), in which a strained Si channel grown onto a relaxed SiGe layer is used. Si–Ge interdiffusion has been studied by many research groups using various SiGe heterostructures and measurement techniques such as SIMS, X-ray diffraction (XRD), Rutherford backscattering spectroscopy (RBS), photoluminescence (PL), and Raman spectroscopy. These studies are listed in Table 12.3. Figure 12.9 shows the scheme of the structure and strain distribution of asymmetrically and symmetrically strained SLs used by Holländer *et al.* (1992) to investigate Si–Ge interdiffusion. The Si–Ge interdiffusion is a mixing process at interfaces between SiGe with different alloy compositions. Therefore, the interdiffusion involves a concentration gradient and occurs under the strain associated with the lattice mismatch. This is significantly different from the isotope self-diffusion described in Section 12.4, where the compositions are uniform without strain and remain unchanged. For the modeling, Si–Ge interdiffusion cannot be treated as dopant diffusion, that is, the dilute-dopant approximation, that the dopant concentration is low enough so that the host lattice does not change, no longer holds. A modeling that takes into account the conservation of lattice sites is found in the work by Hasanuzzaman and Haddara (2008) and Hasanuzzaman *et al.* (2009).

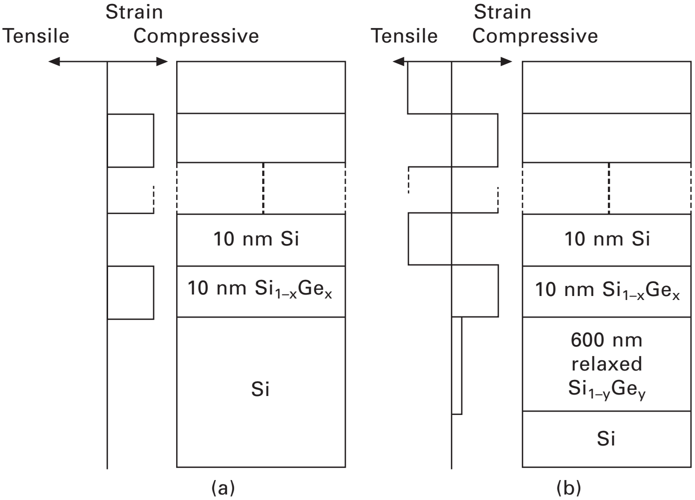
12.5.1 Ge composition dependence

In the studies on Si–Ge interdiffusion with various Ge contents (Table 12.3), the interdiffusivities show an increase with increasing Ge content, as expected from the dependence of self-diffusion in SiGe alloys described in Section 12.4. Figure 12.10 shows the interdiffusivities obtained from the symmetrically strained SLs with different Ge contents (Holländer *et al.*, 1992). In addition, in a heterostructure with a single Ge content, Ge concentration dependence of the interdiffusion is also observed. In a Si/$Si_{0.7}Ge_{0.3}$/Si structure, faster diffusion is observed in the high-concentration part of the Ge profile, leading to an almost rectangular diffusion profile (much flatter and steeper-sided than the Gaussian curves that would be produced by a constant diffusivity), as shown in Fig. 12.11 (Cowern *et al.*, 1994). Furthermore, Fig. 12.12 illustrates the SIMS concentration profiles of Ge in a fully relaxed Ge(250 nm)/Si structure before and after annealing at 750°C (Gavelle *et al.*, 2008a, 2008b). After 3 h, Si has diffused over ~80 nm inside the Ge layer from the Ge/Si interface, whereas the Ge penetration into Si is ~15 nm due to the much higher Si diffusivity in Ge than for the Ge diffusivity in Si (see Fig. 12.7). In addition, the diffused Ge profiles in Fig. 12.12 cross each other

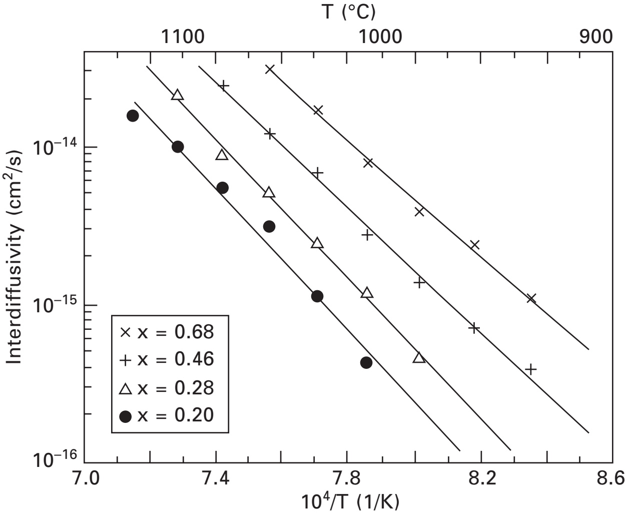
Table 12.3 List of studies on Si–Ge interdiffusion

References	Structure ($Si_{1-x}Ge_x$)[a]	Temperature (°C)	Technique
Chang *et al.*, 1989	SSL Ge/Si on 0.4	640–780	XRD
Iyer and LeGoues, 1989	ASL 0.12/Si on Si	900	RBS
van der Walle *et al.*, 1989	Si/s-0.07, 0.17, 0.33 on Si	775–1010	RBS
Prokes and Wang, 1990	ASL 0.35/Si on Si	700–880	XRD
Prokes *et al.*, 1992	ASL 0.3/Si on Si SSL 0.3/Si on 0.2	900	Raman
Holländer *et al.*, 1992	ASL 0.2, 0.27, 0.45, 0.63, 0.7/Si on Si SSL 0.2, 0.28, 0.46, 0.68/Si on 0.18, 0.23, 0.24, 0.37, respectively	900–1125	RBS
Baribeau, 1993	ASL 0.45/Si on Si	700	XRD
Sunamura *et al.*, 1993, 1994	Si/s-0.16, 0.18 on Si	800–950	PL
Cowern *et al.*, 1994	Si/s-0.1–0.3 on Si	900–1050	SIMS
Zaumseil *et al.*, 1994	Si/s-0.25 on Si	993, 1006	XRD
Cowern *et al.*, 1996	Si/s-0.3 on Si, s-Si/0.3/Si on 0.3	875	SIMS
Takeuchi *et al.*, 2002	Ge (60 nm) on Si	900	SIMS
Aubertine and McIntyre, 2005	s-SL 0.05/0.1, 0.08/0.13, 0.1/0.16, 0.15/0.19, 0.17/0.21 on Si	770–870	XRD
Xia *et al.*, 2006, 2007	s-Si on 0.2, 0.4, 0.45, 0.56 s-Si/0.3, 0.45, 0.56/Si on 0, 0.15, 0.3, 0.56	770–920	SIMS
Ozguven and McIntyre, 2008	s-SL 0.84/0.98 on 0.86	600–700	XRD
Gavelle *et al.*, 2008a, 2008b	Ge (250 nm) on Si	750–900	SIMS

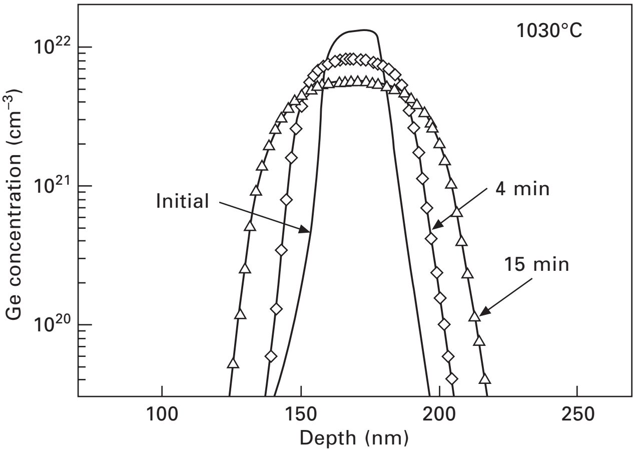
[a] ASL = asymmetrically strained SL; SSL = symmetrically strained SL; s- = strained; 'x/Si on y' refers to a $Si_{1-x}Ge_x$/Si heterostructure on a relaxed $Si_{1-y}Ge_y$ buffer layer.


Strain
Tensile
Compressive
10 nm Si
10 nm $Si_{1-x}Ge_x$
Si
(a)
Strain
Tensile
Compressive
10 nm Si
10 nm $Si_{1-x}Ge_x$
600 nm relaxed $Si_{1-y}Ge_y$
Si
(b)

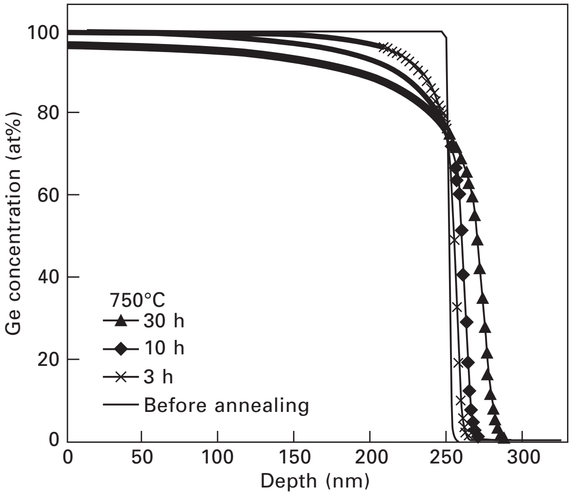
12.9 Scheme of the structure and strain distribution of (a) asymmetrically (ASL) and (b) symmetrically (SSL) strained SLs. ASL is grown on a Si buffer layer, resulting in compressively strained $Si_{1-x}Ge_x$ layers and unstrained Si layers. In SSL that is grown on a relaxed $Si_{1-y}Ge_y$ buffer layer, $Si_{1-x}Ge_x$ layers are compressively strained and Si layers are under tensile strain (when $x > y$). The small amount of residual strain in the buffer layer is sketched in SSL (adapted from Holländer *et al.*, 1992).



12.10 Si–Ge interdiffusivities obtained from symmetrically strained SLs with different Ge contents (adapted from Holländer *et al.*, 1992).



12.11 Ge concentration profiles in Si/Si$_{0.7}$Ge$_{0.3}$/Si structures measured by SIMS before (line) and after annealing (lines + symbols) at 1030°C for 4 and 15 min (adapted from Cowern *et al.*, 1994).



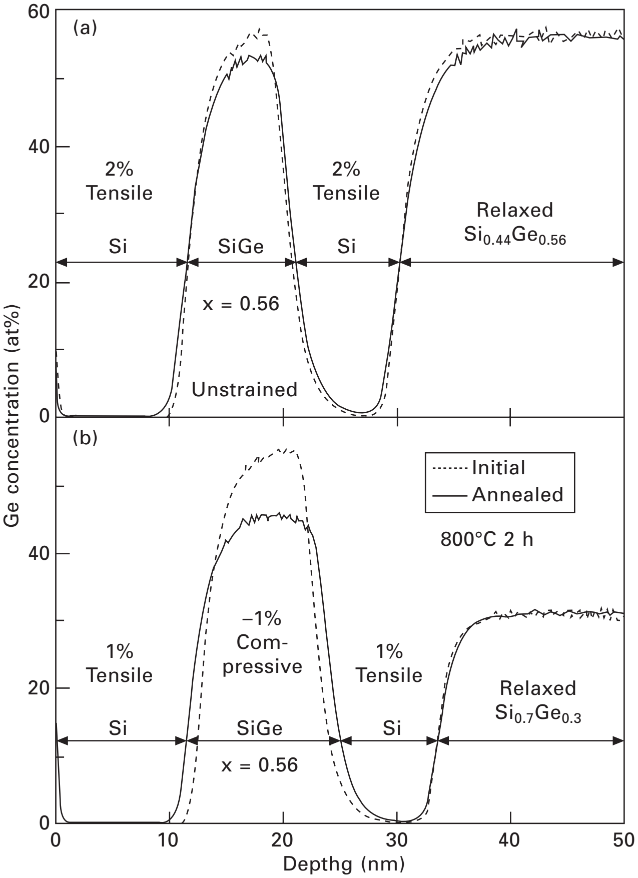
12.12 Ge concentration profiles in Ge/Si structures measured by SIMS before (line) and after annealing (lines + symbols) at 750°C (adapted from Gavelle *et al.*, 2008b).

at the same concentration (~75 at%) near the original Ge/Si interface. This clearly indicates that the Si–Ge interdiffusion is concentration dependent. Otherwise, Fick's laws solution with a single diffusivity would have resulted in an intersection at 50 at%. Similarly, formation of a 50 nm graded Si$_{1-x}$Ge$_x$ layer, with consumption of 30 nm Ge and 20 nm Si, was observed after 900°C annealing of Ge/Si samples (Takeuchi *et al.*, 2002).

12.5.2 Strain effect

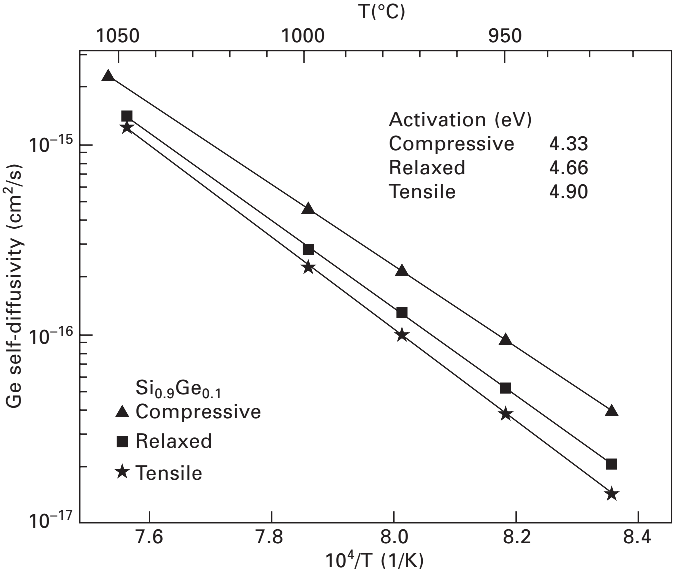
On the other hand, the effect of strain on Si–Ge interdiffusion has not yet been completely established. Several authors found that biaxial compressive strain enhances the interdiffusion (Iyer and LeGoues, 1989; Baribeau, 1993; Cowern *et al.*, 1994, 1996; Xia *et al.*, 2006, 2007). By contrast, some researchers reported that biaxial strain had no detectable influence on interdiffusion rates; no significant difference in the interdiffusivities was found between the asymmetrically and symmetrically strained SLs (Fig. 12.9) (Holländer *et al.*, 1992) and a compressive strain difference of 0.2% had little influence on interdiffusion (Aubertine and McIntyre, 2005). In order to examine the strain effect on interdiffusion, the strain in SLs is varied by using different relaxed $Si_{1-x}Ge_x$ buffer layers. Figure 12.13 shows the Ge profiles of structures with (a) an unstrained $Si_{0.44}Ge_{0.56}$ layer on a relaxed $Si_{0.44}Ge_{0.56}$ buffer layer and (b) a compressively strained $Si_{0.44}Ge_{0.56}$ layer on a relaxed $Si_{0.7}Ge_{0.3}$ buffer layer (Xia *et al.*, 2006, 2007). The authors found significantly enhanced Si–Ge interdiffusion in $Si/Si_{1-x}Ge_x/Si$ heterostructures ($x = 0.3$, 0.45, and 0.56) with $Si_{1-x}Ge_x$ layers under compressive strain of about −1% compared with interdiffusion when the $Si_{1-x}Ge_x$ layers are unstrained. Faster diffusion is observed in the compressively strained layer compared to the unstrained layer and the ratio of the interdiffusivity at the Ge concentration of 50% between the two is about 20 in Fig. 12.13. In addition, Cowern *et al.* (1996) showed that intermixing was faster in $Si/Si_{0.7}Ge_{0.3}/Si$ structures with compressively strained SiGe than it was in $Si/Si_{0.7}Ge_{0.3}/Si/Si_{0.7}Ge_{0.3}$ structures with unstrained SiGe. Concerning the effect on tensile strain, fewer studies have been reported. Xia *et al.* (2006, 2007) concluded that biaxial tensile strain of 1% has no observable effect by comparing tensile strained $Si_{0.7}Ge_{0.3}$ layers on relaxed $Si_{0.44}Ge_{0.56}$ with unstrained $Si_{0.7}Ge_{0.3}$ layers on relaxed $Si_{0.7}Ge_{0.3}$. On the contrary, Prokes *et al.* (1992) reported that interdiffusion in $Si_{0.7}Ge_{0.3}/Si$ SLs was enhanced under tensile strain. Besides, the strain effect of Ge isotope self-diffusion (with no concentration gradient) in $Si_{0.9}Ge_{0.1}$ was examined by Zangenberg *et al.* (2001). To create strained layers, the surface and substrate layers were made with 15% or 5% Ge to produce tensile or compressive strain (±0.21%), respectively. The results in Fig. 12.14 showed a decrease in Ge self-diffusivity and an increase in activation enthalpy upon going from compressive over relaxed to tensile strain. This dependence will be discussed below in Section 12.7.3.

As described above, the effect of strain on Si–Ge interdiffusion has not yet been completely elucidated. In addition, the Ge composition dependence of the activation enthalpy of Si–Ge interdiffusion reported so far shows significantly different trends. These inconsistencies may come from the characteristics of Si–Ge interdiffusion which depends on Ge composition and strain (if the effect exists) that is determined by the Ge concentration



12.13 SIMS Ge profiles of heterostructures with (a) an unstrained $Si_{0.44}Ge_{0.56}$ layer on a relaxed $Si_{0.44}Ge_{0.56}$ buffer layer and (b) a compressively strained $Si_{0.44}Ge_{0.56}$ layer on a relaxed $Si_{0.7}Ge_{0.3}$ buffer layer before and after annealing at 800°C for 2 h. The $Si_{0.44}Ge_{0.56}$ layer in (a) is unstrained and the two surrounding Si layers are under biaxial tensile strain of about 2%. The $Si_{0.44}Ge_{0.56}$ layer in (b) is under biaxial compressive strain of about –1% and the two surrounding Si layers are under about 1% biaxial tensile strain (adapted from Xia *et al.,* 2007).

gradient. In the first place, interdiffusion is concentration dependent and the diffusion profiles cannot be fitted by a single diffusivity (see Figs 12.11 and 12.12). Thus, the obtained interdiffusivity is a certain effective value, which may include some uncertainty. In addition, interdiffusion between the individual layers causes a decrease of the concentration gradient, resulting in strain relaxation. Indeed, strain relaxation due to interdiffusion was reported in several studies listed in Table 12.3. During interdiffusion, therefore, Ge



12.14 Arrhenius plot of Ge self-diffusivities in compressive, relaxed, and strained $Si_{0.90}Ge_{0.10}$ with activation enthalpies of 4.33, 4.66, and 4.90 eV, respectively (adapted from Zangenberg *et al.*, 2001).

(and Si) composition and strain change with time, which further changes the interdiffusion rate. This indicates that the interdiffusivity measured critically depends on the diffusion amount (or diffusion length), which is a function of temperature and time, even if the diffusion starts with the same heterostructures and Ge compositions. Furthermore, the effect of strain may depend on the Ge composition. Biaxial compressive strain enhances vacancy-assisted diffusion in SiGe (Cowern *et al.*, 1994; Pakfar, 2002). On the other hand, interstitial-assisted diffusion is retarded by biaxial compressive strain (Pakfar, 2002). As described in Section 12.4, the contribution of vacancies to self-diffusion in SiGe increases with an increase of Ge composition. This suggests that the effect of compressive strain becomes more significant in heterostructures with higher Ge contents.

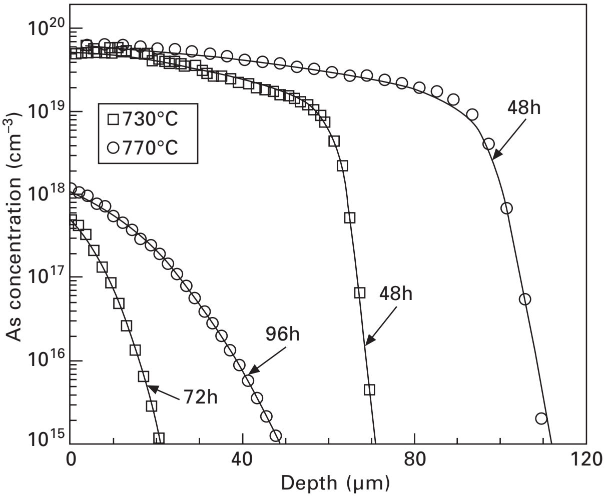
12.6 Dopant diffusion in germanium (Ge)

In this section, the diffusion of Group V (P, As, Sb) and Group III (B) atoms in Ge is described because these atoms are important elements for n- and p-type doping purposes.

12.6.1 Group V atoms

Diffusion experiments of P, As, and Sb in Ge were performed utilizing diluted Ge-dopant alloys as diffusion sources by Bracht and Brotzmann (2006) and Brotzmann and Bracht (2008). Figure 12.15 shows in-diffusion profiles of active As concentration in Ge measured by spreading resistance profiler (SRP) after diffusion annealing at 730 and 770°C. When the As surface concentration is about 10^{18} cm^{-3}, which is below the intrinsic carrier concentration, the profile is described by a complementary error function. For higher As concentration, that is, under extrinsic conditions, a box-shaped As profile is obtained, which is characteristic of concentration-dependent diffusion. Similar profiles were also obtained for P and Sb diffusion. These diffusion profiles of P, As, and Sb in a wide range of diffusion conditions were fitted by the diffusivities which are proportional to the square of the n-type dopant concentration (Bracht and Brotzmann, 2006; Brotzmann and Bracht, 2008). A quadratic dependence of As diffusion on the As concentration was also found in the simultaneous diffusion of Ge and As in As-implanted Ge isotope SLs performed by Naganawa *et al.* (2008a) (see Fig. 12.5).

In the same manner as Ge self-diffusion, n-type dopants in Ge diffuse via the vacancy mechanism (Eq. 12.1) (Frank *et al.*, 1984). In addition, the quadratic dopant concentration dependence of the diffusion shows that the charge difference between A_s and (AV) is two (Gösele, 1988), which means



12.15 Active As concentration profiles in Ge measured by SRP after diffusion annealing at 730 and 770°C. Solid lines represent best fits obtained on the basis of the vacancy mechanism (adapted from Bracht and Brotzmann, 2006).

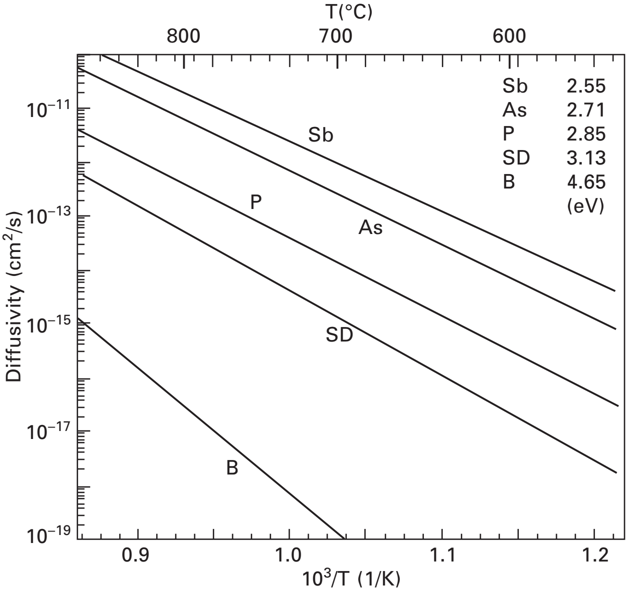
that the charge of (AV) is singly negative. Therefore, n-type dopant diffusion in Ge is described by

$$(AV)^- \Leftrightarrow A_s^+ + V^{2-} \ (A = P, As, Sb) \qquad 12.6$$

and the diffusivities under extrinsic conditions are described by

$$D_A(n) = D_A(n_i)\ (n/n_i)^2 \qquad 12.7$$

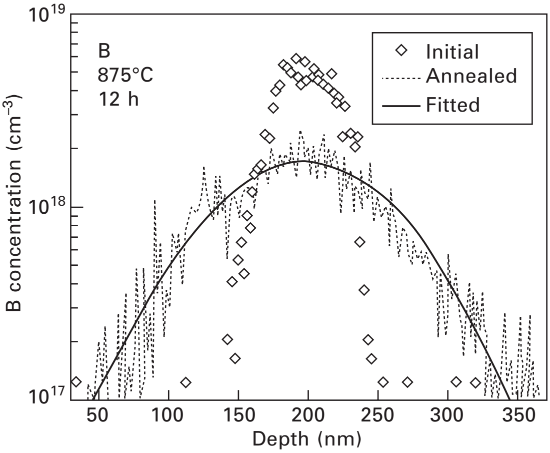
Note that the quadratic dependence does not lead to a doubly negative charge state for the vacancy but leads to a singly negative charge state for (AV). Figure 12.16 shows the intrinsic diffusivities ($D_A(n_i)$) of P, As, and Sb in Ge obtained by Brotzmann and Bracht (2008), and the activation enthalpies and pre-exponential factors of the intrinsic diffusivities are summarized in Table 12.4. The diffusivities of all n-type dopants P, As, and Sb exceed the Ge self-diffusivity and the activation enthalpies of n-type dopants are smaller than that of self-diffusion. This indicates an attractive interaction between the dopants and the vacancies in the vacancy mechanism, as mentioned in Section 12.2. Furthermore, the activation enthalpy of n-type dopant diffusion decreases with increasing size of the dopants, indicating an increasing attractive force between the dopant and the vacancy.



12.16 Arrhenius plot of intrinsic diffusivities of P, As, and Sn in Ge with activation enthalpies (Brotzmann and Bracht, 2008). Intrinsic Ge self-diffusivity (Hüger *et al.*, 2008) and B diffusivity (Uppal *et al.*, 2004) are also shown.

Table 12.4 Intrinsic diffusivities of P, As, Sb and B in Ge

Dopant	*H* (eV)	D_0 (cm^2/s)	Temperature (°C)	Reference
P	2.85	9.1	650–920	Brotzmann and Bracht, 2008
As	2.71	32	640–920	Brotzmann and Bracht, 2008
Sb	2.55	16.7	600–920	Brotzmann and Bracht, 2008
B	4.65	1.97×10^5	800–900	Uppal *et al.*, 2004



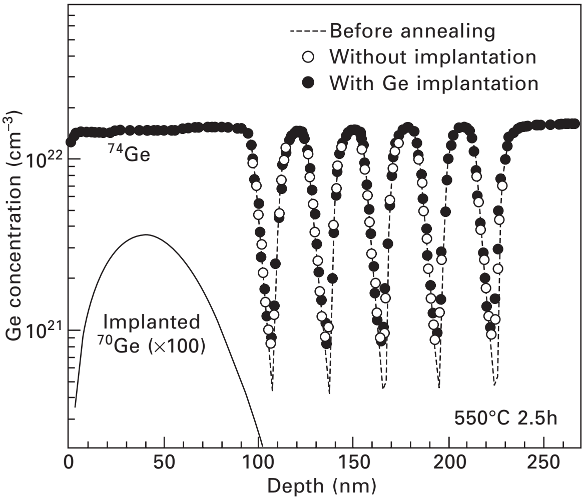
12.17 SIMS profiles of B in epitaxially grown Ge before and after annealing at 875°C for 12 h. The solid line represents a fit to obtain the diffusivity (adapted from Uppal *et al.*, 2004).

12.6.2 Group III atoms

Diffusion of B in Ge was studied using implantation doping and B-doped epitaxial Ge layers by Uppal *et al.* (2004). Figure 12.17 shows the SIMS profiles of B in epitaxially grown Ge before and after annealing at 875°C for 12 h. The obtained intrinsic B diffusivity is plotted in Fig. 12.16 and is listed in Table 12.4. As is clearly seen in the figure, B diffusion in Ge is much slower than n-type dopant diffusion and Ge self-diffusion. This is quite different from B diffusion in Si, which is faster than n-type dopant diffusion and Si self-diffusion (Frank *et al.*, 1984). In Si, B is known to diffuse by means of self-interstitials via the kick-out mechanism (Eq. 12.2) (Frank *et al.*, 1984; Uematsu, 1997). In Ge, however, the contribution of self-interstitials is quite minor, as described in Section 12.3. The significantly low B diffusivity with the high activation enthalpy of 4.65 eV reflects the B diffusion in Ge via interstitial-mediated diffusion, or the kick-out mechanism, as suggested by Brotzmann and Bracht (2008). For other group III dopants, Al, Ga, and In, the diffusion mechanisms are not yet clear due to the spread in the data reported for the activation enthalpies (see Stolwijk and Bracht, 1998), although diffusion via the vacancy mechanism is theoretically suggested (Chroneos *et al.*, 2008a). Quite recently, experimental diffusion profiles of In in Ge were described based on the vacancy mechanism (Kube *et al.*, 2009c). To our knowledge, B seems to be the only dopant in Ge which diffuses via the interstitial-mediated mechanism.

12.6.3 Transient enhanced diffusion

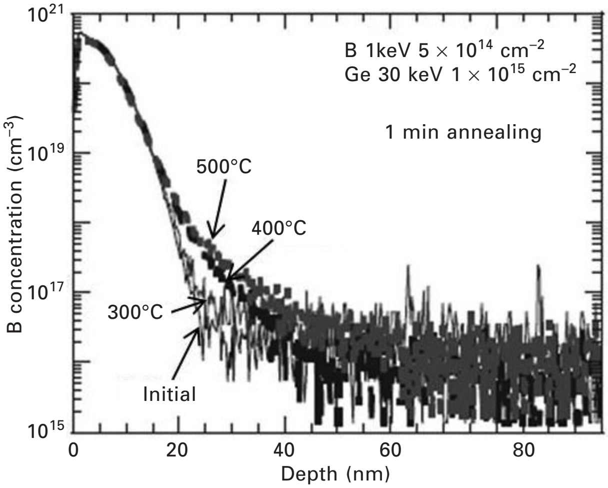
A striking difference between the diffusion in Ge and that in Si is seen in TED. Regarding TED in Si, {311} self-interstitial clusters produced by implantation damage are the sources of supersaturated Si self-interstitials, which enhance dopant diffusion (Stolk *et al.*, 1997). In Ge, TED is much less significant than in Si, as will be described below. The B diffusivities obtained from the epitaxially grown Ge (Fig. 12.17) and those from the B-implanted Ge were found to lie within experimental errors (Uppal *et al.*, 2004). In addition, Ge self-diffusion and As diffusion were found to be in thermal equilibrium with no time dependence for the As-implanted Ge, whose dose would be high enough to induce TED in Si, because the thermal Ge self-diffusivity and As diffusivity described above well reproduced the SIMS profiles in Fig. 12.5 (Naganawa *et al.*, 2008a). Furthermore, the same SLs were implanted with Ge at 90 keV and 2×10^{14} cm^{-2} and annealed at 550°C for 2.5 h (Fig. 12.18). The Ge self-diffusion in the Ge-implanted



12.18 SIMS depth profiles of ^{74}Ge in the ^{nat}Ge/^{70}Ge isotope SLs implanted with ^{70}Ge at 90 keV, 2×10^{14} cm^{-2} and without implantation after annealing at 550°C for 2.5 h. The broken line represents the profile before annealing. The calculated ^{70}Ge profile is also shown (adapted from Uematsu *et al.*, 2009).

SLs showed no significant difference from that without Ge implantation (Uematsu *et al.*, 2009). In addition, the Ge self-diffusivity to fit the data agreed well with the reported thermal intrinsic diffusivities (Fig. 12.4). These results indicate that TED is not observed under the experimental conditions employed in these studies.

Quite recently, however, an indication of TED was observed in annealing at low temperatures and for short times. Figure 12.19 shows the SIMS profiles of B implanted at 1 keV with a dose of 5×10^{14} cm^{-2} into pre-amorphized Ge (30 keV, 1×10^{15} cm^{-2}) before and after annealing at 300, 400, and 500°C for 1 min (Simoen *et al.*, 2008). Fast B diffusion was observed in the tail region for short anneals and this tail diffusion is much faster than that estimated from the B intrinsic diffusivity (Table 12.4). The observation of TED at low temperatures and for short annealing times is consistent with the recent TEM results by Koffel *et al.* (2009). They observed the extended defects of interstitial type upon the solid-phase epitaxial regrowth of amorphous layers created by ion implantation in Ge. The defects are more difficult to observe than in Si, being smaller in size and density. In addition, these defects disappear very quickly during annealing. Concerning direct observation of the Ge self-interstitial contributions, enhanced Ge intermixing of Ge isotope SLs was observed under proton irradiation (Schneider *et al.*, 2008). Moreover, the TED in Fig. 12.19 was not observed without Ge pre-amorphization, indicating that the interstitials generated during the B implantation itself are not sufficient to generate TED (Simoen *et al.*, 2008).



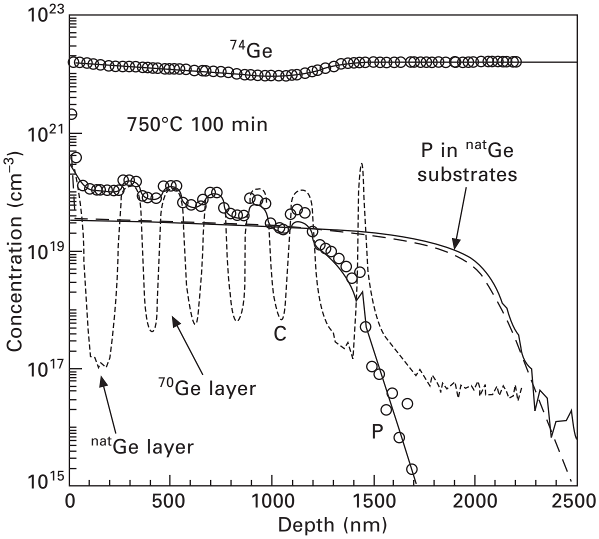
12.19 SIMS profiles of B implanted at 1 keV, 5×10^{14} cm^{-2} to pre-amorphized Ge (30 keV, 1×10^{15} cm^{-2}) before and after annealing at 300, 400, and 500°C for 1 min (adapted from Simoen *et al.*, 2008).

This indicates that much higher implantation doses would be necessary to induce TED in Ge than in Si. In summary, TED in Ge due to the implantation damage, if present, is much less significant than that in Si. This is closely related to a minor contribution of self-interstitials in Ge because TED occurs due to supersaturated self-interstitials.

12.6.4 Dopant activation

In order to obtain lower contact resistance for Ge devices, dopant activation has been investigated for B (Chui *et al.*, 2003; Satta *et al.*, 2005; Mirabella *et al.*, 2008), Ga (Impellizzeri *et al.*, 2009), and n-type dopants (Chui *et al.*, 2003, 2005; Satta *et al.*, 2006a; Koike *et al.*, 2008). Dopants with a concentration above the thermal solid solubility (Satta *et al.*, 2006b) become inactive in thermal equilibrium conditions. However, higher active dopant concentration than the solubility can be obtained in non-equilibrium conditions, such as ion-implanted Ge. Implantation of B to pre-amorphized Ge and subsequent annealing result in active dopant concentrations of a few 10^{20} cm^{-3} (Satta *et al.*, 2005), which is well above the solid solubility of B (6×10^{18} cm^{-3}). In addition, B is a slow diffuser, and thus B is suitable for shallow junction formation. On the other hand, it still remains difficult to get equivalently high active n-type dopant concentrations in Ge. The deactivation mechanism of n-type dopants is not yet understood, although the dopant precipitation (Satta *et al.*, 2006a) and the formation of neutral donor-vacancy complexes (such as As_2V and As_4V for As) (Brotzmann *et al.*, 2008; Chroneos *et al.*, 2008b) seem to be the cause. Besides, for another p-type dopant, Ga has a higher solid solubility than B, but higher Ga diffusivity is a drawback.

Finally, the effect of carbon (C) on dopant behavior is described. As shown in Fig. 12.20, the trapping of P within the C-doped layers and the reduction of P diffusion by the presence of C were observed (Brotzmann *et al.*, 2008). The self-diffusion was not affected by the presence of C even though the diffusion of donors was reduced. The dopant trapping and diffusion retardation were strongest for Sb and similar for P and As. This effect of C is attributed to the formation of immobile neutral C-vacancy–donor complexes (Chroneos *et al.*, 2008b). From the viewpoint of process applications, the suppression of donor diffusion by the incorporation of C is helpful in realizing shallow junction formation. This effect of C on the suppression of impurity diffusion is also known for Si. The population of excess interstitials induced by ion implantation was found to be strongly reduced by incorporating C in Si (Stolk *et al.*, 1997). This demonstrates that C acts as traps for self-interstitials, thereby suppressing the interstitial-mediated impurity diffusion in Si.



12.20 Concentration profiles of P (lower symbols), ^{74}Ge (upper symbols), and C (short dashed line) measured by SIMS after diffusion annealing of the Ge isotope SLs at 750°C for 100 min. The experimental P profile (thin solid line) obtained in natural Ge substrates under the identical condition is shown for comparison. The upper and lower solid lines and thin wide dashed lines are the simulation results. In the Ge isotope SLs, ^{70}Ge layers were doped with C, whereas the C concentration of the adjacent ^{nat}Ge layers was below the detection limit of SIMS. C doping of the Ge isotope SLs reduces P diffusion compared to P diffusion in high-purity ^{nat}Ge substrates. The C distribution remained unchanged after diffusion annealing (adapted from Brotzmann *et al.*, 2008).

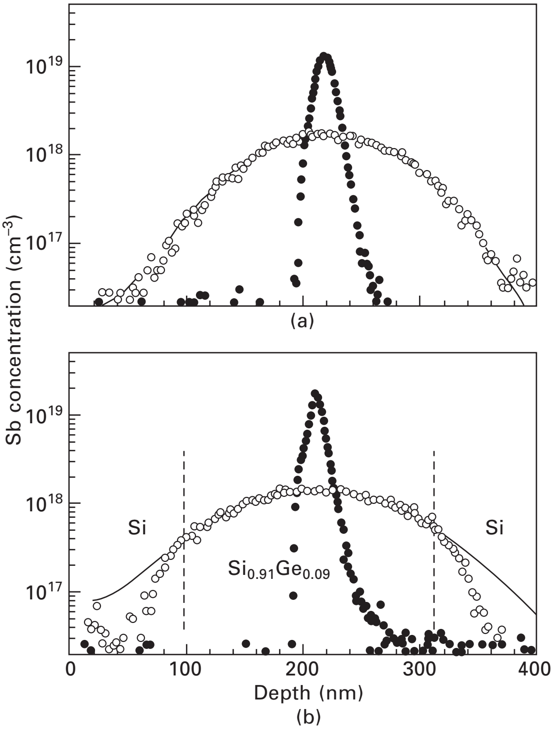
12.7 Dopant diffusion in silicon–germanium (SiGe) alloys

In Si, B and P diffuse via the interstitial-mediated mechanism (Eq. 12.2), while Sb diffuses via the vacancy mechanism (Eq. 12.1) and As diffuses via both the interstitial-mediated and the vacancy mechanisms (Frank *et al.*, 1984; Uematsu, 1997). In contrast, P, As, and Sb in Ge diffuse via the vacancy mechanism, while B in Ge is interstitial mediated, as described in Section 12.6. This indicates that there may be a changeover from interstitial-mediated to vacancy-mediated diffusion at some Ge content.

12.7.1 Group V atoms

First, Sb diffusion in SiGe is described because Sb simply diffuses via the vacancy mechanism. Diffusion of Sb in relaxed $Si_{1-x}Ge_x$ alloy layers grown by MBE was studied as a function of the Ge composition between 0 and 0.5 (Nylandsted-Larsen and Kringhøj, 1996). The layer consisted of an undoped graded buffer layer on which a layer of SiGe with a constant Ge content was grown, including a buried Sb doped layer. The Sb diffusivity increased with the Ge alloy content, and the extracted activation enthalpies for diffusion decreased with increasing Ge. The effect of strain on the Sb diffusion in Ge was investigated by comparing the diffusion in compressively strained and relaxed $Si_{0.91}Ge_{0.09}$ (Kringhøj *et al.*, 1996). The strained $Si_{0.91}Ge_{0.09}$ layer with an Sb-doped layer was sandwiched by Si layers. As shown in Fig. 12.21, the Sb diffusion was enhanced by the compressive strain. These results of the dependencies on Ge content and strain are consistent with vacancy-mediated Sb diffusion in both Si and Ge.

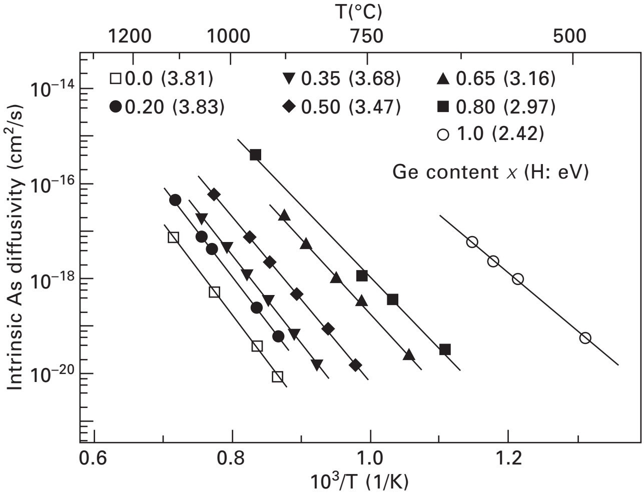
Diffusion of As in relaxed $Si_{1-x}Ge_x$ alloy epilayers was studied as a function of Ge composition x with 0, 0.20, 0.35, 0.50, 0.65, 0.8, and 1 (Laitinen *et al.*, 2003). The diffusion of As was performed by means of a radiotracer



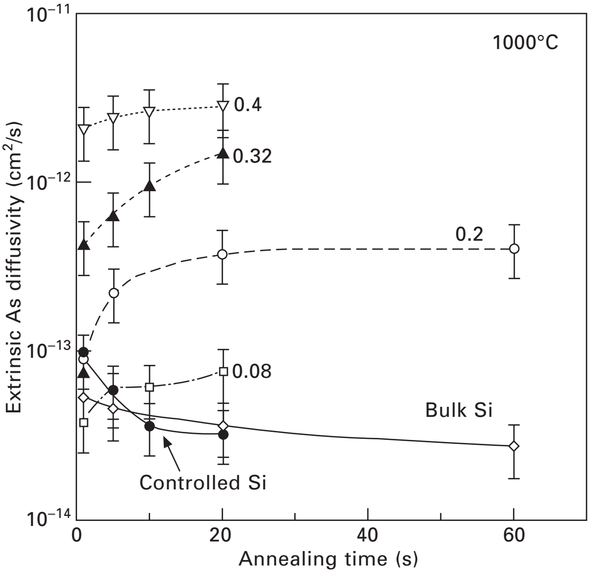
12.21 Sb profiles in (a) relaxed and (b) compressively strained $Si_{0.91}Ge_{0.09}$ measured by SIMS before and after annealing at 1028°C for 30 min. Solid lines represent the fitting results. The boundaries between Si and SiGe are indicated by the dashed lines (adapted from Kringhøj *et al.*, 1996).

technique employing ^{73}As. The As diffusivity was found to increase with the Ge alloy content, and the activation enthalpies decreased with increasing Ge content, as shown in Fig. 12.22. The same trend was observed in ion-implanted As in relaxed $Si_{1-x}Ge_x$ alloys with x = 0, 0.08, 0.2, 0.32, and 0.4 under thermal equilibrium conditions (long annealing times) by Eguchi *et al.* (2002, 2004). In their experiments, transient enhanced and retarded diffusion of As in SiGe was observed depending on the Ge content for short annealing times (Fig. 12.23). For the Ge content of 0.08, the As diffusivity appeared to be constant with time, and the TED, which is evident in Si (x = 0), was not observed. For the Ge content of 0.2 or above, the As diffusion in $Si_{1-x}Ge_x$ was observed to be retarded for short annealing times, compared to the thermal diffusivities extracted for longer annealing times. These results suggest that the vacancy contribution in As diffusion increases with an increase of Ge content, assuming that excess Si interstitials generated by implantation reduce the vacancy concentration in SiGe and that this excess decreases with the Ge content.

Diffusion of P in relaxed $Si_{1-x}Ge_x$ alloys with a buried P-doped layer was investigated as a function of the Ge composition between 0 and 0.4 (Zangenberg *et al.*, 2003). The P diffusivity appeared to increase with the Ge alloy content; however, the enhancement is not as high as that for As diffusion. The same trend was observed by Eguchi *et al.* (2002). Diffusion in compressively strained $Si_{1-x}Ge_x$ (x = 0–0.22) with a buried P-doped layer was investigated by Christensen *et al.* (2001). No significant dependence of



12.22 Arrhenius plot of intrinsic As diffusivities in relaxed $Si_{1-x}Ge_x$ layers for x = 0, 0.20, 0.35, 0.50, 0.65, 0.8, and 1 with activation enthalpies H (adapted from Laitinen *et al.*, 2003).

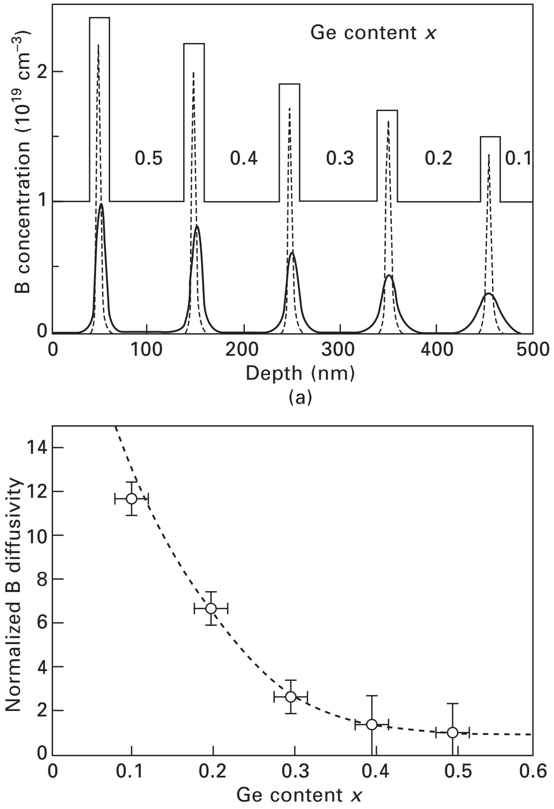


12.23 Time dependence of the time-averaged extrinsic As diffusivities in $Si_{1-x}Ge_x$ alloys with $x = 0, 0.08, 0.2, 0.32$, and 0.4 that were implanted with As at 15 keV, 3×10^{15} cm^{-2}. The extrinsic As diffusivities were evaluated at the As concentration of 2×10^{20} cm^{-3} (adapted from Eguchi *et al.*, 2004).

the P diffusivity on the Ge content was observed, suggesting that composition and strain effects compensate for each other. From the oxidation of the same SiGe, P diffusion in $Si_{1-x}Ge_x$ ($x \lesssim 0.2$) was found to be primarily governed by self-interstitials. These results seem to be related to mixed contributions of interstitials and vacancies; however, less is clear compared to Sb and As due to fewer reported results.

12.7.2 Group III atoms

For the diffusion of Group III atoms, B is the only dopant that has been considerably studied. Diffusion of B in compressively strained $Si_{1-x}Ge_x$ layers was investigated as a function of the Ge content (Moriya *et al.*, 1993). A strained multilayer structure contained five $Si_{1-x}Ge_x$ layers with Ge content in the range of 0.1 to 0.5, separated by 50 nm of undoped Si space layers. Thin B-doped regions (5 nm) were introduced in the center of each $Si_{1-x}Ge_x$ layer, as illustrated in Fig. 12.24. B diffusion in the compressively strained SiGe was reduced with increasing Ge content, through the authors' explanation in terms of band-gap narrowing appears to be incorrect, as pointed out by Cowern *et al.* (1994). In order to separate the effects of strain and Ge content, Kuo *et al.* (1995a) investigated B diffusion in $Si_{1-x}Ge_x$ layers with

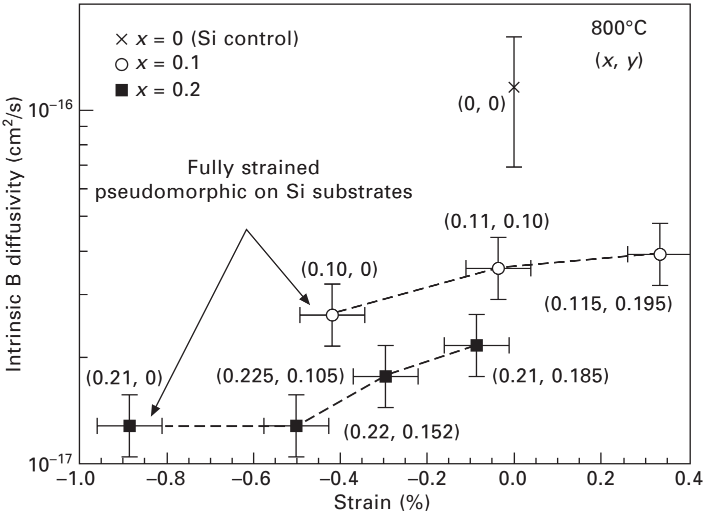


(a)

(b)

12.24 (a) B concentration in a strained multilayer structure with different Ge content in each layer measured by SIMS before (dashed line) and after (solid line) annealing at 975°C for 70 s. (b) Intrinsic B diffusivity normalized to its value for $x = 0.5$ as a function of Ge content (adapted from Moriya *et al.*, 1993).

a constant Ge content ($x \simeq 0.1$ and 0.2) grown on various relaxed $Si_{1-y}Ge_y$ ($0 \leq y \leq 0.2$) substrates. In this structure, a B-doped region (20 nm) was introduced in the center of the $Si_{1-x}Ge_x$ layer (60 nm) and the amount of strain in the $Si_{1-x}Ge_x$ layer (the diffusion layer) was varied independently of the Ge content x, with different Ge content y (Fig. 12.25). Boron diffusion primarily depends on Ge content in the $Si_{1-x}Ge_x$ layers, and did not show a strong dependence on biaxial strain. This indicates that the strain effect does not account for the slower B diffusion in $Si_{1-x}Ge_x$ observed by Moriya *et al.* (1993). Boron diffusion in compressively strained $Si_{0.88}Ge_{0.12}$ and $Si_{0.76}Ge_{0.24}$ and tensile strained $Si_{0.88}Ge_{0.12}$ was also investigated by Zangenberg *et al.* (2003). Compressive strain was found to decrease the B diffusivity and tensile strain to increase it. In addition, the results for the oxidation (interstitial

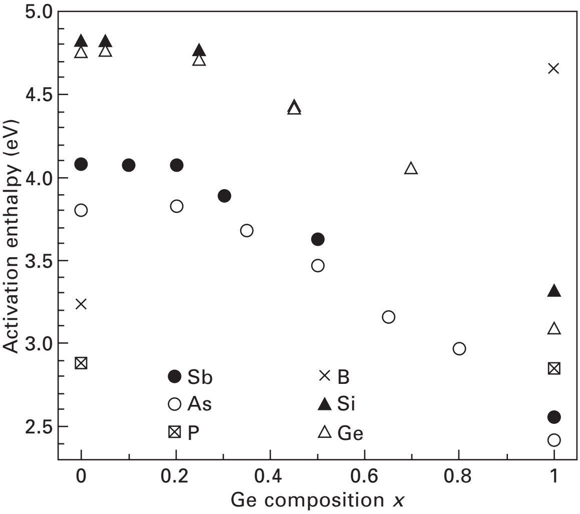


12.25 Intrinsic B diffusivity at 800°C as a function of strain in $Si_{1-x}Ge_x$ grown on $Si_{1-y}Ge_y$ substrates. The (*x*, *y*) label for each data point represents the Ge contents in the diffusion layer and the substrate, respectively. Positive strain represents biaxial tension and negative strain represents biaxial compression (adapted from Kuo *et al.*, 1995a).

injection) of compressively strained $Si_{0.82}Ge_{0.18}$ (Kuo *et al.*, 1995b) and for the oxidation and nitridation (vacancy injection) of compressively strained $Si_{0.8}Ge_{0.2}$ (Fang *et al.*, 1996) revealed that B atoms in $Si_{1-x}Ge_x$ ($x \simeq 0.2$) primarily diffuse by interstitials. These results are consistent with interstitial-mediated B diffusion in both Si and Ge.

12.7.3 Activation enthalpy of diffusion

The dependence of the activation enthalpies of dopant diffusion on the Ge content of $Si_{1-x}Ge_x$ is illustrated in Fig. 12.26, including that of Si and Ge diffusion (self-diffusion). The values of Sb diffusion are from Nylandsted-Larsen and Kringhøj (1996), those of As from Laitinen *et al.* (2003), and P and B in Si from Naganawa *et al.* (2008b). The values of Sb, P, and B in Ge and those of self-diffusion (Si and Ge) are listed in Tables 12.4 and 12.2, respectively. The composition dependence of the activation enthalpies of Sb and As diffusion shows an upward bowing. Similar composition dependence is predicted for the stability of phosphorus-vacancy and arsenic-vacancy pairs in SiGe (Chroneos *et al.*, 2008c). In addition, the self-diffusion shows the same trend of an upward bowing, as already described in Section 12.4.2. Moreover, the Ge self-diffusion in $Si_{0.9}Ge_{0.1}$ showed a decrease in self-diffusivity and an increase in activation enthalpy upon going from compressive over relaxed



12.26 Composition dependence of the activation enthalpies of dopant diffusion in $Si_{1-x}Ge_x$ alloys. Sb: Nylandsted-Larsen and Kringhøj (1996); As: Laitinen *et al.* (2003); P and B in Si: Naganawa *et al.* (2008b); Sb, P, and B in Ge: Table 12.4. Self-diffusion (Si and Ge) data are in Table 12.2.

to tensile strain (Fig. 12.14), indicating the contributions of both interstitials and vacancies. Thus, the upward bowing of the activation enthalpy suggests that the contribution of vacancies to self-diffusion in SiGe alloys remains close to that in Si up to the Ge content of about 0.25 and rapidly increases above 0.5. This is consistent with the conclusion by Laitinen *et al.* (2003) that both interstitials and vacancies contribute as diffusion vehicles in the composition range $0 \leq x \leq 0.35$ and that the vacancy mechanism dominates diffusion in the range $0.35 \leq x \leq 1$.

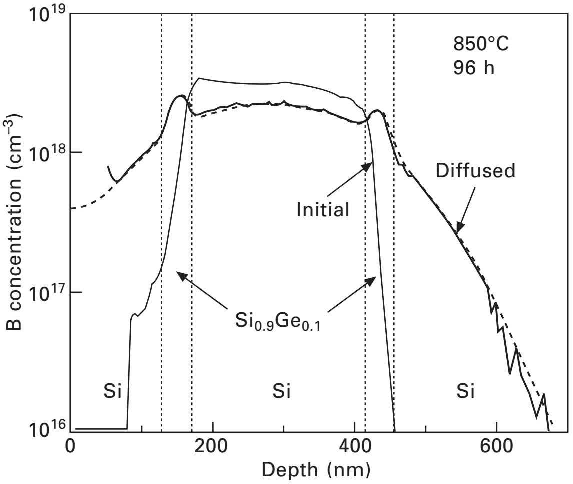
12.8 Dopant segregation

When dopants diffuse across the interface of SiGe heterostructures, dopant atoms often exhibit a prominent segregation behavior among the layers. Hu *et al.* (1991) performed diffusion experiments of B, P, and As in $Si/Si_{0.9}Ge_{0.1}$(70 nm)/Si heterostructures and found that B tends to segregate into Ge-rich layers, while P and As tend to segregate away from it. The segregation coefficients (the ratio between the thermal equilibrium concentration of dopants in SiGe and that in Si) of B, P, and As through the $Si_{0.9}Ge_{0.1}/Si$ interface at 950°C were 1.35, 0.76, and 0.75, respectively. Boron segregation through the $Si_{0.8}Ge_{0.2}/Si$ interface was investigated by Moriya *et al.* (1995) and its temperature dependence was obtained (~2.5 at 900°C, 2.0 at 1050°C, and 1.7 at 1150°C). Boron segregation through the $Si_{0.8}Ge_{0.2}/Si$ interface at

850°C was 3 (Fang *et al.*, 1996). In the study of Lever *et al.* (1998), B was introduced to an epitaxial Si layer confined between two $Si_{1-x}Ge_x$ layers and the diffused B profiles showed two concentration peaks in these confining $Si_{1-x}Ge_x$ layers (Fig. 12.27). They fitted the profiles by assuming a formation of Ge–B pairs; however, it appears to be highly speculative and the existence of the pair has not been experimentally detected, as pointed out by Chen *et al.* (1999). Rather, Chen *et al.* (1999) fitted the same B profiles by assuming that the dopant segregation behavior is attributed to the solubility difference of the ionized shallow acceptor species in the different layer materials, which is determined by a chemical effect, a Fermi-level effect, and an effect of the junction carrier concentrations. In their simulation, smaller intrinsic B diffusivities with increasing Ge content were used, which is consistent with the experimental results described in Section 12.7.2.

12.9 Conclusion and future trends

In self-diffusion in Ge, doubly negatively charged vacancies are the dominant point defects and the contribution of self-interstitials is very minor, which is quite different from the diffusion in Si. Group V atoms in Ge diffuse mainly via doubly negatively charged vacancies, but B in Ge shows very slow diffusion due to the minor self-interstitial contribution in Ge. Even when dopants are introduced by ion implantation, TED in Ge due to the damage, if present, is much less significant than that in Si. Self-diffusion in SiGe



12.27 B concentration profiles in an epitaxial Si layer confined between two $Si_{0.9}Ge_{0.1}$ layers measured by SIMS before and after annealing at 850°C for 96 h. The dashed line represents the calculated results (adapted from Lever *et al.*, 1998).

alloys and Si–Ge interdiffusion become faster with increasing Ge content. The Ge composition dependence of self-diffusion and dopant diffusion in SiGe with the effects of strain suggests that both interstitials and vacancies contribute as diffusion vehicles up to a Ge content of about 0.3 and that the vacancy mechanism dominates the diffusion above 0.3. For dopant segregation, B tends to segregate into Ge-rich layers, while P and As tend to segregate away from it.

Diffusion mechanisms in Ge are almost established; however, realizing abrupt junctions for n-type dopants remains a challenge due to the low donor activation level and to the higher diffusivities compared to that of B. TED in SiGe will become an issue as the device size shrinks, although TED in Ge is much less significant than that in Si. In order to understand and to predict the TED in SiGe, the changes in the contributions of self-interstitials and vacancies with Ge content should be more precisely determined. The effects of oxidation (or nitridation) and of strain on self-diffusion in SiGe alloys using isotope heterostructures and the diffusion of various dopants will help us to fully understand the diffusion mechanisms in SiGe. In addition, simultaneous observations of self-diffusion and dopant diffusion in SiGe can determine the charge states of point defects in SiGe, which determine the diffusion in the extrinsic conditions. Furthermore, the knowledge will facilitate the further understanding of the characteristics of Si–Ge interdiffusion which depends on Ge composition and strain.

With the scaling down of material structures, the effect of interface on diffusion may become serious. Such an influence has been observed in the interface between Si and Si oxide, in which SiO molecules generated at the interface diffuse into the oxide to enhance the self-diffusion and B diffusion (Fukatsu *et al*., 2003; Uematsu *et al*., 2004a, 2004b, 2004c). SiGe materials are often covered with Si oxide or nitride films, resulting in the interfaces with oxide or nitride near the surface. In MOS devices, interfaces with high-*k* gate dielectrics exist. Moreover, there is an interface with oxide in the bulk if SiGe-on-insulator (SGOI) or GOI substrates are used. Point defects may be generated at these interfaces upon annealing, which could enhance the diffusion in SiGe and oxide. This effect would be more serious with the regions of interest being closer to the interfaces. Extensive studies on the interface effect are expected by means of isotope heterostructures.

12.10 Sources of further information and advice

This chapter focuses on an update of the recent literature, taking into account recent considerable progress in this area thanks to the availability of stable isotopes. Overviews on the early results can be found in Frank *et al*. (1984) and Stolwijk and Bracht (1998). In addition, this chapter is devoted to the diffusion of self-atoms and of Group III and V dopants. The references on metal diffusion, which occurs via the direct interstitial mechanism (Frank *et al*. 1984), can also be found in these references. The latest information on the subjects of this chapter can be found in the websites of the research groups listed below.

Isotope engineering: University of California at Berkeley, USA (E. E. Haller) http://www.mse.berkeley.edu/faculty/Haller/Haller.html

Growth of isotope heterostructures: University of Aarhus, Denmark (A. Nylandsted Larsen) http://www.phys.au.dk/~anl

Experiments on dopant diffusion: IMEC, Belgium (E. Simoen and C. Claeys) http://www2.imec.be/imec_com/scaling-driven-nanoelectronics.php

Diffusion experiments, modeling and simulation: University of Münster, Germany (H. Bracht)
http://www.uni-muenster.de/Physik.MP/Bracht/index.shtml

Growth of isotope heterostructures, diffusion experiments, modeling and simulation: University of Keio, Japan (M. Uematsu and K. M. Itoh)
http://www.appi.keio.ac.jp/Itoh_group

Experiments on Si–Ge interdiffusion: Stanford University, USA (P. C. McIntyre) http://www.stanford.edu/group/mcintyre/index.html

12.11 References

Asen-Palmer M, Bartkowski K, Gmelin E, Cardona M, Zhernov A P, Inyushkin A V, Taldenkov A, Ozhogin V I, Itoh K M and Haller E E (1997), ‘Thermal conductivity of germanium crystals with different isotopic compositions’, *Phys Rev B*, 56, 9431–9447.

Aubertine D B and McIntyre P C (2005), ‘Influence of Ge concentration and compressive biaxial stress on interdiffusion in Si-rich SiGe alloy heterostructures’, *J Appl Phys*, 97, 013531 1–10.

Aubry J C, Tyliszczak T, Hitchcock A P, Baribeau J-M and Jackman T E (1999), ‘First-shell bond lengths in $Si_{1-x}Ge_x$ crystalline alloys’, *Phys Rev B*, 59, 12872–12883.

Baribeau J-M (1993), ‘X-ray reflectometry study of interdiffusion in Si/Ge heterostructures’, *J Appl Phys*, 74, 3805–3810.

Bracht H (2007), ‘Self- and foreign-atom diffusion in semiconductor isotope heterostructures. I. Continuum theoretical calculations’, *Phys Rev B*, 75, 035210 1–16.

Bracht H and Brotzmann S (2006), ‘Atomic transport in germanium and the mechanism of arsenic diffusion’, *Mater Sci Semicond Process*, 9, 471–476.

Bracht H, Haller E E and Clark-Phelps R (1998), ‘Silicon self-diffusion in isotope heterostructures’, *Phys Rev Lett*, 81, 393–396.

Brotzmann S and Bracht H (2008), ‘Intrinsic and extrinsic diffusion of phosphorus, arsenic, and antimony in germanium’, *J Appl Phys*, 103, 033508 1–7.

Brotzmann S, Bracht H, Lundsgaard-Hansen J, Nylandsted-Larsen A, Simoen E, Haller E E, Christensen J S and Werner P (2008), ‘Diffusion and defect reactions between donors, C, and vacancies in Ge. I. Experimental results’, *Phys Rev B*, 77, 235207 1–13.

Capinski W S, Maris H J, Bauser E, Silier I, Asen-Palmer M, Ruf T, Cardona M and Gmelin E (1997), ‘Thermal conductivity of isotopically enriched Si’, *Appl Phys Lett*, 71, 2109–2111.

Chang S J, Wang K L, Bowman Jr. R C and Adams P M (1989), 'Interdiffusion in a symmetrically strained Ge/Si superlattice', *Appl Phys Lett*, 54, 1253–1255.

Chen C-H, Gösele U M and Tan T Y (1999), 'Dopant diffusion and segregation in semiconductor heterostructures: Part 2. B in Ge_xSi_{1-x}/Si structures', *Appl Phys A*, 68, 19–24.

Christensen J S, Kuznetsov A Y, Radamson H H and Svensson B G (2001), 'Phosphorus diffusion in $Si_{1-x}Ge_x$', *Defect Diffus Forum*, 194–199, 709–716.

Chroneos A, Bracht H, Grimes R W and Uberuaga B P (2008a), 'Vacancy-mediated dopant diffusion activation enthalpies for germanium', *Appl Phys Lett*, 92, 172103 1–3.

Chroneos A, Grimes R W, Uberuaga B P and Bracht H (2008b), 'Diffusion and defect reactions between donors, C, and vacancies in Ge. II. Atomistic calculations of related complexes', *Phys Rev B*, 77, 235208 1–7.

Chroneos A, Bracht H, Jiang C, Uberuaga B P and Grimes R W (2008c), 'Nonlinear stability of E centers in $Si_{1-x}Ge_x$: Electronic structure calculations', *Phys Rev B*, 78, 195201 1–7.

Chui C O, Gopalakrishnan K, Griffin P B, Plummer J D and Saraswat K C (2003), 'Activation and diffusion studies of ion-implanted *p* and *n* dopants in germanium', *Appl Phys Lett*, 83, 3275–3277.

Chui C O, Kulig L, Moran J, Tsai W and Saraswat K C (2005), 'Germanium *n*-type shallow junction activation dependences', *Appl Phys Lett*, 87, 091909 1–3.

Cowern N E B, Zalm P C, van der Sluis P, Gravesteijn D J and de Boer W B (1994), 'Diffusion in strained Si(Ge)', *Phys Rev Lett*, 72, 2585–2588.

Cowern N E B, Kersten W J, de Kruif R C M, van Berkum J G M, de Boer W B, Gravesteijn D J and Bulle-Liewma C W T (1996), 'Interdiffusion mechanisms in coherently strained SiGe multilayers', *Electrochem Soc Proc*, 96–4, 195–209.

Eguchi S, Hoyt J L, Leitz C W and Fitzgerald E A (2002), 'Comparison of arsenic and phosphorus diffusion behavior in silicon–germanium alloys', *Appl Phys Lett*, 80, 1743–1745.

Eguchi S, Chleirigh C N, Olubuyide O O and Hoyt J L (2004), 'Germanium-concentration dependence of arsenic diffusion in silicon germanium alloys', *Appl Phys Lett*, 84, 368–370.

Fang T T, Fang W T C, Griffin P B and Plummer J D (1996), 'Calculation of the fractional interstitial component of boron diffusion and segregation coefficient of boron in $Si_{0.8}Ge_{0.2}$', *Appl Phys Lett*, 68, 791–793.

Frank W, Gösele U, Mehrer H and Seeger A (1984), 'Diffusion in silicon and germanium', in Murch G E and Nowick A S, *Diffusion in Crystalline Solids*, New York, Academic Press, 63–142.

Fuchs H D, Walukiewicz W, Haller E E, Dondl W, Schorer R, Abstreiter G, Rudnev A I, Tikhomirov A V and Ozhogin V I (1995), 'Germanium $^{70}Ge/^{74}Ge$ isotope heterostructures: An approach to self-difffusion studies', *Phys Rev B*, 51, 16817–16821.

Fukatsu S, Takahashi T, Itoh K M, Uematsu M, Fujiwara A, Kageshima H, Takahashi Y, Shiraishi K and Gösele U (2003), 'Effect of the Si/SiO_2 interface on self-diffusion of Si in semiconductor-grade SiO_2', *Appl Phys Lett*, 83, 3897–3899.

Gavelle M, Bazizi E M, Scheid E, Armand C, Fazzini P F, Marcelot O, Campidelli Y, Halimaoui A and Cristiano F (2008a), 'Study of silicon–germanium interdiffusion from pure germanium deposited layers', *Mater Sci Eng B*, 154–155, 110–113.

Gavelle M, Bazizi E M, Scheid E, Fazzini P F, Cristiano F, Armand C, Lerch W, Paul S, Campidelli Y and Halimaoui A (2008b), 'Detailed investigation of Ge–Si interdiffusion in the full range of $Si_{1-x}Ge_x (0 \leq x \leq 1)$ composition', *J Appl Phys*, 104, 113524 1–7.

Gösele U M (1988), 'Fast diffusion in semiconductors', *Annu Rev Mater Sci*, 18, 257–282.

Hasanuzzaman M and Haddara Y M (2008), 'Modeling silicon–germanium interdiffusion by the vacancy exchange and interstitial mechanisms', *J Mater Sci: Mater Electron*, 19, 569–576.

Hasanuzzaman M, Haddara Y M and Knights A P (2009), 'Modeling germanium diffusion in $Si_{1-x}Ge_x$/Si superlattice structures', *J Appl Phys*, 105, 043504 1–5.

Holländer B, Butz R and Mantl S (1992), 'Interdiffusion and thermally induced strain relaxation in strained $Si_{1-x}Ge_x$/Si superlattices', *Phys Rev B*, 46**,** 6975–6981.

Hu S M, Ahlgren D C, Ronsheim P A and Chu J O (1991), 'Experimental study of diffusion and segregation in a Si-(Ge_xSi_{1-x}) heterostructure', *Phys Rev Lett*, 67, 1450–1453.

Hüger E, Tietze U, Lott D, Bracht H, Bougeard D, Haller E E and Schmidt H (2008), 'Self-diffusion in germanium isotope multilayers at low temperatures', *Appl Phys Lett*, 93, 162104 1–3.

Impellizzeri G, Mirabella S, Irrera A, Grimaldi M G and Napolitani E (2009), 'Ga-implantation in Ge: Electrical activation and clustering', *J Appl Phys*, 106, 013518 1–6.

Itoh K M (2005), 'An all-silicon linear chain NMR quantum computer', *Solid State Commun*, 133, 747–752.

Itoh K M, Hansen W L, Haller E E, Farmer J W, Ozhogin V I, Rudnev A and Tikhomirov A (1993), 'High purity isotopically enriched ^{70}Ge and ^{74}Ge single crystals: Isotope separation, growth, and properties', *J Mater Res*, 8, 1341–1347.

Iyer S S and LeGoues F K (1989), 'Thermal relaxation of pseudomorphic Si–Ge superlattices by enhanced diffusion and dislocation multiplication', *J Appl Phys*, 65, 4693–4698.

Karaiskaj D, Thewalt M L W, Ruf T, Cardona M, Pohl H-J, Deviatych G G, Sennikov P G and Riemann H (2001), 'Photoluminescence of isotopically purified silicon: How sharp are bound exciton transitions?', *Phys Rev Lett*, 86, 6010–6013.

Koffel S, Cherkashin N, Houdellier F, Hytch M J, Benassayag G, Scheiblin P and Claverie A (2009), 'End of range defects in Ge', *J Appl Phys*, 105, 126110 1–3.

Koike M, Kamata Y, Ino T, Hagishima D, Tatsumura K, Koyama M and Nishiyama A (2008), 'Diffusion and activation of *n*-type dopants in germanium', *J Appl Phys*, 104, 023523 1–5.

Kringhøj P, Nylandsted-Larsen A and Shirayev S Y (1996), 'Diffusion of Sb in strained and relaxed Si and SiGe', *Phys Rev Lett*, 76, 3372–3375.

Kube R, Bracht H, Lundsgaard-Hansen J, Nylandsted-Larsen A, Haller E E, Paul S and Lerch W (2009a), 'Simultaneous diffusion of Si and Ge in isotopically controlled $Si_{1-x}Ge_x$ heterostructures', *Mater Sci Semicond Process*, doi: 10.1016/j.mssp.2008.07.005.

Kube R, Bracht H, Lundsgaard-Hansen J, Nylandsted-Larsen A, Haller E E, Paul S and Lerch W (2009b), 'Self-diffusion in $Si_{1-x}Ge_x$ isotope heterostructures and its composition dependence', *Abstract of the 25th International Conference of Defects in Semiconductors (St Petersburg, July 2009)*, Mon-1.44po.

Kube R, Bracht H, Chroneos A, Posselt M and Schmidt B (2009c), 'Intrinsic and extrinsic diffusion of indium in germanium', *J Appl Phys*, 106, 063534 1–7.

Kuo P, Hoyt J L, Gibbons J F, Turner J E and Lefforge D (1995a), 'Effects of strain on boron diffusion in Si and $Si_{1-x}Ge_x$', *Appl Phys Lett*, 66, 580–582.

Kuo P, Hoyt J L, Gibbons J F, Turner J E and Lefforge D (1995b), 'Effects of Si thermal oxidation on B diffusion in Si and strained $Si_{1-x}Ge_x$ layers', *Appl Phys Lett*, 67, 706–708.

Ladd T D, Goldman J R, Yamaguchi F, Yamamoto Y, Abe E and Itoh K M (2002), 'All-silicon quantum computer', *Phys. Rev. Lett.* 89, 017901 1–4.

Laitinen P, Riihimäki I, Räisänen J and The ISOLDE Collaboration (2003), 'Arsenic diffusion in relaxed $Si_{1-x}Ge_x$', *Phys Rev B*, 68, 155209 1–6.

Lever R F, Bonar J M and Willoughby A F W (1998), 'Boron diffusion across silicon–silicon germanium boundaries', *J Appl Phys*, 83, 1988–1994.

McVay G L and DuCharme A R (1974), 'Diffusion of Ge in SiGe alloys', *Phys Rev B*, 9, 627–631.

Mirabella S, Impellizzeri G, Piro A M, Bruno E and Grimaldi M G (2008), 'Activation and carrier mobility in high fluence B implanted germanium', *Appl Phys Lett*, 92, 251909 1–3.

Morita K, Itoh K M, Muto J, Mizoguchi K, Usami N, Shiraki Y and Haller E E (2000), 'Growth and characterization of $^{70}Ge_n/^{74}Ge_n$ isotope superlattices', *Thin Solid Films*, 369, 405–408.

Moriya N, Feldman L C, Luftman H S, King C A, Bevk J and Freer B (1993), 'Boron diffusion in strained $Si_{1-x}Ge_x$ epitaxial layers', *Phys Rev Lett*, 71, 883–886.

Moriya N, Feldman L C, Downey S W, King C A and Emerson A B (1995), 'Interfacial segregation in strained heterostructures: Boron in $Si_{0.8}Ge_{0.2}$/Si', *Phys Rev Lett*, 75, 1981–1983.

Naganawa M, Shimizu Y, Uematsu M, Itoh K M, Sawano K, Shiraki Y and Haller E E (2008a), 'Charge states of vacancies in germanium investigated by simultaneous observation of germanium self-diffusion and arsenic diffusion', *Appl Phys Lett*, 93, 191905 1–3.

Naganawa M, Kawamura Y, Shimizu Y, Uematsu M, Itoh K M, Ito H, Nakamura M, Ishikawa H and Ohji Y (2008b), 'Accurate determination of the intrinsic diffusivities of boron, phosphorus, and arsenic in silicon: The influence of SiO_2 films', *Jpn J Appl Phys*, 47, 6205–6207.

Nylandsted-Larsen A and Kringhøj P (1996), 'Diffusion of Sb in relaxed $Si_{1-x}Ge_x$', *Appl Phys Lett*, 68, 2684–2686.

Ozguven N and McIntyre P C (2008), 'Silicon–germanium interdiffusion in high-germanium-content epitaxial heterostructures', *Appl Phys Lett*, 92, 181907 1–3.

Pakfar A (2002), 'Dopant diffusion in SiGe: modeling stress and Ge chemical effects', *Mater Sci Eng B*, 89, 225–228.

Prokes S M and Wang K L (1990), 'Interdiffusion measurements in asymmetrically strained SiGe/Si superlattices', *Appl Phys Lett*, 56, 2628–2630.

Prokes S M, Glembocki O J and Godbey D J (1992), 'Stress and its effect on the interdiffusion in $Si_{1-x}Ge_x$/Si superlattices', *Appl Phys Lett*, 60, 1087–1089.

Ruf T, Henn R W, Asen-Palmer M, Gmelin E, Cardona M, Pohl H-J, Devyatych G G and Sennikov P G (2000), 'Thermal conductivity of isotopically enriched silicon', *Solid State Commun*, 115, 243–247.

Satta A, Simoen E, Clarysse T, Janssens T, Benedetti A, De Jaeger B, Meuris M and Vandervorst W (2005), 'Diffusion, activation, and recrystallization of boron implanted in preamorphized and crystalline germanium', *Appl Phys Lett*, 87, 172109 1–3.

Satta A, Simoen E, Duffy R, Janssens T, Clarysse T, Benedetti A, Meuris M and Vandervorst W (2006a), 'Diffusion, activation, and regrowth behavior of high dose P implants in Ge', *Appl Phys Lett*, 88, 162118 1–3.

Satta A, Simoen E, Janssens T, Clarysse T, De Jaeger B, Benedetti A, Hoflijk I, Brijs B, Meuris M and Vandervorst W (2006b), 'Shallow junction ion implantation in Ge and associated defect control', *J Electrochem Soc*, 153, G229–G233.

Schneider S, Bracht H, Petersen M C, Lundsgaard-Hansen J and Nylandsted-Larsen A (2008), 'Proton irradiation of germanium isotope multilayer structures at elevated temperatures', *J Appl Phys*, 103, 033517 1–5.

Shimizu Y and Itoh K M (2006), 'Growth and characterization of short-period silicon isotope superlattices', *Thin Solid Films*, 508, 160–162.

Shimizu Y, Uematsu M and Itoh K M (2007), 'Experimental evidence of the vacancy-mediated silicon self-diffusion in single-crystalline silicon', *Phys Rev Lett*, 98, 095901 1–4.

Shimizu Y, Uematsu M, Itoh K M, Takano A, Sawano K and Shiraki Y (2009), 'Behaviors of neutral and charged silicon self-interstitials during transient enhanced diffusion in silicon investigated by isotope superlattices', *J Appl Phys*, 105, 013504 1–6.

Shockley W and Moll J L (1960), 'Solubility of flaws in heavily-doped semiconductors', *Phys Rev*, 119, 1480–1482.

Silveira E, Dondl W, Abstreiter G and Haller E E (1997), 'Ge self-diffusion in isotopic $(^{70}\mathrm{Ge})_m/(^{74}\mathrm{Ge})_n$ superlattices: A Raman study', *Phys Rev B*, 56, 2062–2069.

Silvestri H H, Bracht H, Lundsgaard-Hansen J, Nylandsted-Larsen A and Haller E E (2006), 'Diffusion of silicon in crystalline germanium', *Semicond Sci Technol*, 21, 758–762.

Simoen E, Brouwers G, Satta A, David M-L, Pailloux F, Parmentier B, Clarysse T, Goossens J, Vandervorst W and Meuris M (2008), 'Shallow boron implantations in Ge and the role of the pre-amorphization depth', *Mater Sci Semicond Process*, doi: 10.1016/j.mssp.2008.09.006.

Spitzer J, Ruf T, Cardona M, Dondl W, Schorer R, Abstreiter G and Haller E E (1994), 'Raman scattering of optical phonons in isotopic $^{70}(\mathrm{Ge})_n/^{74}(\mathrm{Ge})_n$ superlattices', *Phys Rev Lett*, 72, 1565–1568.

Stolk P A, Gossmann H-J, Eaglesham D J, Jacobson D C, Rafferty C S, Gilmer G H, Jaraiz M, Poate J M, Luftman H S and Haynes T E (1997), 'Physical mechanisms of transient enhanced dopant diffusion in ion-implanted silicon', *J Appl Phys*, 81, 6031–6050.

Stolwijk N A and Bracht H (1998), 'Diffusion in silicon, germanium and their alloys', *Landolt-Börnstein New Series*, III/33A, New York, Springer.

Strohm A, Voss T, Frank W, Laitinen P and Räisänen J (2002), 'Self-diffusion of ^{71}Ge and ^{31}Si in Si–Ge alloys', *Z Metallkunde*, 93, 737–744.

Sunamura H, Fukatsu S, Usami N and Shiraki Y (1993), 'Luminescence study on interdiffusion in strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$/Si single quantum wells grown by molecular beam epitaxy', *Appl Phys Lett*, 63, 1651–1653.

Sunamura H, Fukatsu S, Usami N and Shiraki Y (1994), 'Optical detection of interdiffusion in strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$/Si quantum well structures', *Jpn J Appl Phys*, 33, 2344–2347.

Takeuchi H, Ranade P, Subramanian V and King T-J (2002), 'Observation of dopant-mediated intermixing at Ge/Si interface', *Appl Phys Lett*, 80, 3706–3708.

Uematsu M (1997), 'Simulation of boron, phosphorus, and arsenic diffusion in silicon based on an integrated diffusion model, and the anomalous phosphorus diffusion mechanism', *J Appl Phys*, 82, 2228–2246.

Uematsu M, Kageshima H, Takahashi Y, Fukatsu S, Itoh K M, Shiraishi K and Gösele U (2004a), 'Modeling of Si self-diffusion in SiO_2: Effect of the Si/SiO_2 interface including time-dependent diffusivity', *Appl Phys Lett*, 84, 876–878.

Uematsu M, Kageshima H, Takahashi Y, Fukatsu S, Itoh K M and Shiraishi K (2004b), 'Correlated diffusion of silicon and boron in thermally grown SiO_2', *Appl Phys Lett*, 85, 221–223.

Uematsu M, Kageshima H, Takahashi Y, Fukatsu S, Itoh K M and Shiraishi K (2004c), 'Simulation of correlated diffusion of Si and B in thermally grown SiO_2', *J Appl Phys*, 96, 5513–5519.

Uematsu M, Naganawa M, Shimizu Y, Itoh K M, Sawano K, Shiraki Y and Haller E E (2009), 'Probing the behaviors of point defects in silicon and germanium using isotope superlattices', *Electrochem Soc Trans*, 25, 3, 51–54.

Uppal S, Willoughby A F W, Bonar J M, Cowern N E B, Grasby T, Morris R J H and Dowsett M G (2004), 'Diffusion of boron in germanium at 800–900°C', *J Appl Phys*, 96, 1376–1380.

van de Walle G F A, van Ijzendoorn L J, van Gorkum A A, van den Heuvel R A, Theunissen A M L and Gravesteijn D J (1989), 'Germanium diffusion and strain relaxation in Si/$Si_{1-x}Ge_x$/Si structures', *Thin Solid Films*, 183, 183–190.

Werner M, Mehrer H and Hochheimer H D (1985), 'Effect of hydrostatic pressure, temperature, and doping on self-diffusion in germanium', *Phys Rev B*, 32, 3930–3937.

Xia G, Oluwamuyiwa O, Olubuyide O, Hoyt J L and Canonico M (2006), 'Strain dependence of Si–Ge interdiffusion in epitaxial Si/$Si_{1-y}Ge_y$/Si heterostructures on relaxed $Si_{1-x}Ge_x$ substrates', *Appl Phys Lett*, 88, 013507 1–3.

Xia G, Hoyt J L and Canonico M (2007), 'Si–Ge interdiffusion in strained Si/strained SiGe heterostructures and implications for enhanced mobility metal-oxide-semiconductor field-effect transistors', *J Appl Phys*, 101, 044901 1–11.

Zangenberg N R, Lundsgaard-Hansen J, Fage-Pederson J and Nylandsted-Larsen A (2001), 'Ge self-diffusion in epitaxial $Si_{1-x}Ge_x$ layers', *Phys Rev Lett*, 87, 1259011 1–4.

Zangenberg N R, Fage-Pederson J, Lundsgaard-Hansen J and Nylandsted-Larsen A (2003), 'Boron and phosphorus diffusion in strained and relaxed Si and SiGe', *J Appl Phys*, 94, 3883–3890.

Zaumseil P, Jagdhold U and Krüger D (1994), 'X-ray *in situ* observation of relaxation and diffusion processes in $Si_{1-x}Ge_x$ layers on silicon substrates', *J Appl Phys*, 76, 2191–2196.

13

Dislocations and other strain-induced defects in silicon–germanium (SiGe) nanostructures

M. L. LEE, Yale University, USA

Abstract: Strain in SiGe leads to a host of useful effects that can be exploited in device applications, including modulation of band structure, creation of band offsets and quantum confinement, and changes in growth mode. However, another natural consequence of introducing strain into SiGe alloys is the occurrence of plastic deformation via the creation of dislocations and other extended defects. Such defects can adversely affect device performance by a number of mechanisms that vary with the device's principle of operation. The high nucleation barrier for surface dislocation half-loop nucleation combined with the low concentration of fixed dislocation sources dominated many of the early observations of extended defects in SiGe on Si(001). This chapter analyzes further consequences of metastability as well as cases where the barrier to dislocation introduction is lower.

Key words: quantum dot, dislocation, stacking fault, 111, 110.

13.1 Introduction and background

Over past two decades, the vast majority of research on SiGe and its device applications has focused on strained layer epitaxy on (001) substrates. Strain in SiGe leads to a host of useful effects that can be exploited in devices, including modulation of band structure, creation of band offsets and quantum confinement, and changes in growth mode. However, another natural consequence of introducing strain into SiGe alloys is the occurrence of plastic deformation via the creation of dislocations and other extended defects. This chapter reviews some basic aspects of extended defects in SiGe thin films and nanostructures without providing an exhaustive review (Hull and Bean 1992; Fitzgerald, 1991) of the subject. Special attention is given to the influence of defects on SiGe nanostructures such as quantum dots (QDs) and to recent research where issues relating to partial dislocations, stacking faults, and microtwins have become prominent. Finally, some future areas of importance are suggested.

13.1.1 Line defects

Most dislocations in SiGe are 60° dislocations on a {111}<110> slip system. For very high mismatch SiGe films, sessile edge dislocations with *a*/2<110> burgers vector can also frequently be observed due to reactions between glissile 60° dislocations (Marshall *et al.*, 2005). Shockley partial dislocations with a burgers vector of *a*/6<112> on {111} slip planes can also be observed in SiGe, either as a result of 60° dislocation dissociation or from direct surface nucleation. Since the burgers vector *a*/6<112> does not correspond to a full lattice displacement, glide of partials creates stacking faults by definition, as discussed below.

13.1.2 Planar defects

The vast majority of work on SiGe has been on Si(001) where stacking fault formation is rare, particularly for low mismatch epilayers. For higher mismatch layers, stacking faults are observed, and are thought to arise from an initial three-dimensional (3D) growth mode followed by coalescence (Kamins *et al.*, 1997; LeGoues *et al.*, 1990). When 3D growth is avoided, stacking faults are generally not observed for SiGe on Si(001), even when the lattice mismatch is large (Luan *et al.*, 1999). Interestingly, reversing the sign of strain, as is the case for tensile strained Si on relaxed SiGe (001), precipitates significant stacking fault nucleation (Kimura *et al.*, 2006). The reason for the differences in behavior between tensile and compressive films in terms of stacking fault generation is discussed in significant detail in Section 13.3 of this chapter. In addition, SiGe grown on other low-index surface orientations of Si, such as (111) and (110), is often plagued with stacking faults (Hull *et al.*, 1991a; Lee *et al.*, 2006). In this chapter, it is shown that the stacking faults observed in tensile strained Si on relaxed SiGe(001) and in compressive SiGe on Si(110) and (111) are intrinsically intertwined.

13.1.3 Effects of extended defects on silicon–germanium (SiGe) devices

While it is generally accepted that defects have deleterious effects on most types of electronic and optoelectronic devices, it is important to consider the device's principle of operation in order to identify the physical mechanism by which defects harm performance. SiGe has been explored for *minority* carrier devices such as solar cells and for *majority* carrier devices such as field-effect transistors (FETs). Minority carrier devices rely upon gradients in minority carrier concentration and their resulting diffusion currents, while majority carrier devices rely on drift currents that flow due to external bias.

Due to this fundamental difference in operation, minority and majority carrier devices may be affected differently by defects.

Defects in SiGe generally lead to electrically active midgap states, as can be observed by cathodoluminescence and electron beam induced current (EBIC) imaging (Higgs *et al.*, 1992; Higgs and Kittler, 1993). Whether the midgap states result from the dislocation core itself or from impurities that are attracted to the strain field surrounding the dislocation remains unknown (Radzimski *et al.*, 1992). In addition, the possibility of the dislocation core atoms reconstructing means that there is not necessarily a simple relationship between dislocation length and the concentration of midgap states (Duesbery *et al.*, 1991).

The effect of dislocations and other defects on minority carrier devices is most readily understood by analyzing their effect on the trap-assisted minority carrier lifetime of electrons and holes, $\tau_{n,p}$, defined as (Sze, 1981)

$$\tau_{\mathrm{n,p}} = \frac{1}{c_{\mathrm{n,p}} N_{\mathrm{T}}}$$

where $c_{\mathrm{n,p}}$ is the capture coefficients of electrons and holes and N_{T} is the trap concentration. A low value of $\tau_{n,p}$ translates directly into a large reverse saturation current I_{o} for a p–n junction, since:(Sze, 1981)

$$I_{\mathrm{o}} = Agn_{\mathrm{i}}^2 \left(\sqrt{\frac{D_{\mathrm{n}}}{\tau_{\mathrm{n}}}} \cdot \frac{1}{N_{\mathrm{A}}} + \sqrt{\frac{D_{\mathrm{p}}}{\tau_{\mathrm{p}}}} \cdot \frac{1}{N_{\mathrm{D}}} \right)$$

where A is the junction area, q is the electronic charge, n_{i} is the intrinsic carrier concentration, $D_{\mathrm{n,p}}$ is the diffusivity of electrons and holes, respectively, N_{A} is the concentration of acceptor atoms on the p-side, and N_{D} is the concentration of donors on the n-side. Detectivity, the primary figure of merit for a photodetector, scales inversely with I_{o} (Sze, 1981). For a solar cell, the open-circuit voltage V_{oc} is (Sze, 1981)

$$V_{\mathrm{oc}} = \frac{kT}{q} \ln \left(\frac{I_{\mathrm{sc}}}{I_{\mathrm{o}}} + 1 \right)$$

where I_{sc} is the short-circuit current of the solar cell under illumination. For a bipolar junction transistor, the common-emitter current gain β is linearly proportional to $\tau_{\mathrm{n,p}}$ (Sze, 1981).

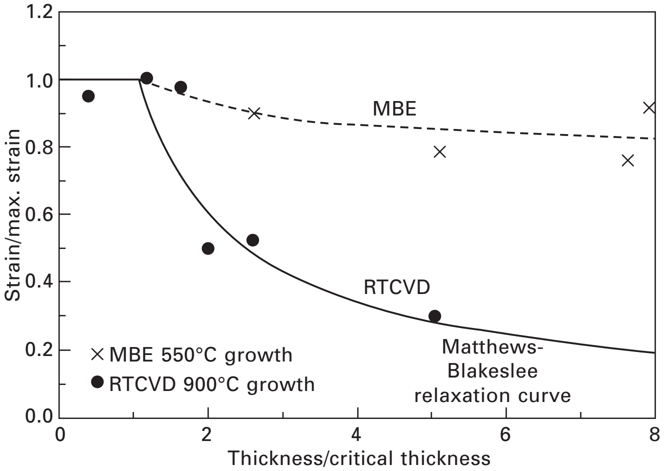
SiGe majority carrier devices are affected by defects through several distinct mechanisms. Dislocations and other defects can scatter carriers and reduce their mobility, leading to a corresponding drop in drive current for an FET. Although, the dislocation scattering effects are much more significant at low temperature due to the dominance of phonon scattering at room temperature (Schäffler *et al.*, 1992), Lee and Fitzgerald (2004) showed that a sufficiently dislocated strained Si layer could exhibit significantly degraded electron mobility. Another effect of dislocations on strained Si FETs was postulated by Fiorenza *et al.* (2004) based on the fact that the region surrounding a dislocation can act as a fast-diffusion 'pipe' for dopants. Fiorenza theorized that a misfit dislocation running directly from source-to-drain in a strained Si MOSFET would result in a shunting path, increasing the off-state current. A final consequence of the midgap states accompanied by dislocations and other defects is the possibility of type conversion, which can affect both minority- and majority-carrier devices. Grillot *et al.* (1996) found that lightly n-doped SiGe graded buffers were weakly p-type when grown at 650°C, but could be converted back to lightly n-type by annealing at T ≥ 800°C, presumably due to elimination of point defects.

13.2 Historical overview

13.2.1 Metastable nature of SiGe layers on Si(001)

Among the earliest observations regarding strained SiGe films and nanostructures on Si(001) was their extreme metastability against dislocation nucleation. For example, Bean *et al.* (1984) observed that SiGe films on Si(001) could be grown 10× thicker than the critical thickness predicted by Matthews and Blakeslee (1974). In fact, the sluggishness of dislocation nucleation led them to question Matthews' analysis of the equilibrium between dislocation energy and elastic strain energy in thin films (People and Bean, 1985). However, other investigators eventually realized that the typical molecular beam epitaxy (MBE) growth temperature of 550°C combined with the high perfection of Si substrates were the main reasons for the extraordinary metastability seen in those structures. Higher temperature growths at 900°C via rapid thermal chemical vapor deposition (RTCVD) showed that strained layers exhibit near-equilibrium behavior; given enough thermal energy, strained films relax via dislocations according to equilibrium theory, as seen in Fig. 13.1 (Green *et al.*, 1991).

While the work of Green *et al.* (1991) showed that high-temperature growth allowed strained SiGe to approach equilibrium, their work said nothing about the sources of dislocations. Early attempts to analyze the energy of dislocation nucleation via the surface half-loop mechanism found values of hundreds of electronvolts for low-mismatch materials, effectively ruling out this mechanism as a source (Fitzgerald, 1991). Fitzgerald *et al.* (1988) studied the effect of substrate patterning on dislocation nucleation in III–V materials, discovering that strained layers grown on small patterns (tens to hundreds of microns on a side) could be grown with extremely low densities of dislocations. The misfit dislocation density dropped linearly as a function of size, indicating that fixed dislocation sources such as surface



13.1 Effect of growth temperature on relaxation in SiGe on Si(001). Adapted from Green *et al.* (1991).

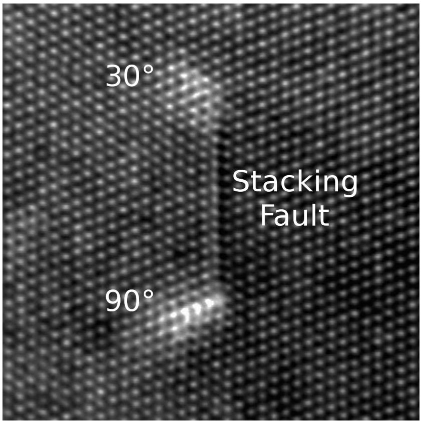
contamination and damage were the primary nucleation sites. Extending these studies to patterned area growth of SiGe on Si(001), Fitzgerald *et al.* (1990) found that the metastability of low-mismatch SiGe films persisted even at high growth temperatures of 900°C. Importantly, this discovery showed that the kinetic barrier to dislocation nucleation in low-mismatch materials is large even in regimes with very high dislocation velocity.

13.2.2 Stranski–Krastanov growth of SiGe nanostructures

The subject of SiGe on Si(001) nanostructures or quantum dots is covered in Chapters 10 and 11, but the connection between high dislocation nucleation energy and 3D growth bears mention here. While observations of Stranski–Krastanov (S–K) growth in highly mismatched materials systems had already been made several years earlier, the work of Eaglesham and Cerullo (1990) showed that Ge islands were dislocation-free over a wide range of sizes, despite being much taller than the critical thickness. Strain relaxation in highly mismatched materials can be treated as a competition between the plastic relaxation allowed by dislocation nucleation and the elastic relaxation enabled by roughening (Tersoff and Legoues, 1994). Eaglesham and Cerullo (1990) noted that the high energy of dislocation half-loop introduction coupled with the inherently small size of the Ge dots were among the necessary factors for coherent islanding; in order to realize dislocation-free dots, roughening must take place before dislocation nucleation. Of course, once the Ge dot is large enough, dislocations are eventually introduced to relieve the high strain energy (Eaglesham and Cerullo, 1990).

13.2.3 Effect of surface orientation on defects in SiGe

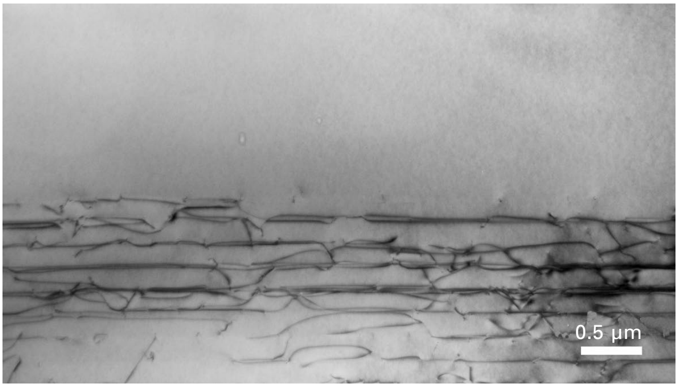
Although the vast majority of work on SiGe and its defects has been concerned with growth on Si(001), Hull *et al.* (1991b) were among the first to examine the growth of SiGe on Si(110) and (111). Interestingly, the emergence of mobility engineering techniques to enhance CMOS performance has generated renewed interest in the influence of surface orientation on carrier mobility, particularly Si(110) surfaces for p-FETs (Yang *et al.*, 2006) and Ge(111) and (110) surfaces for ultra-scaled n-FETs (Low *et al.*, 2004; Takagi, 2003). Surface orientation greatly affects carrier mobility, but also strongly influences the dislocation morphology in mismatched epitaxial semiconductor films. Citing the early work on misfit dislocations of Marée *et al.* (1987), Hull pointed out several important differences between growth on (001) and other low-index surfaces: first, the critical thickness for growth on Si(111) should be higher than on Si(001) due to the lower resolved shear on the inclined {111} glide planes. Conversely, the critical thickness for growth on Si(110) should be lower than on Si(001) due to the higher resolved shear on {111} glide planes. In addition, relaxation should be anisotropic on (110) surfaces, because two of the {111} glide planes are normal to the surface, resulting in zero resolved shear. Finally, for compressively strained films on (111) and (110), a sufficiently high strain can drive separation of 60° total dislocations into 30° and 90° Shockley partial dislocations separated by a stacking fault, as seen in Fig. 13.2. (Hull *et al.*, 1991b).



13.2 High-resolution transmission electron micrograph (TEM) of narrowly dissociated 60° dislocation in SiGe showing stacking fault ribbon between the two partial dislocation cores. Picture area ~112 Å × 112 Å (Lee, 2005).

13.2.4 Low threading dislocation density compositionally graded buffers

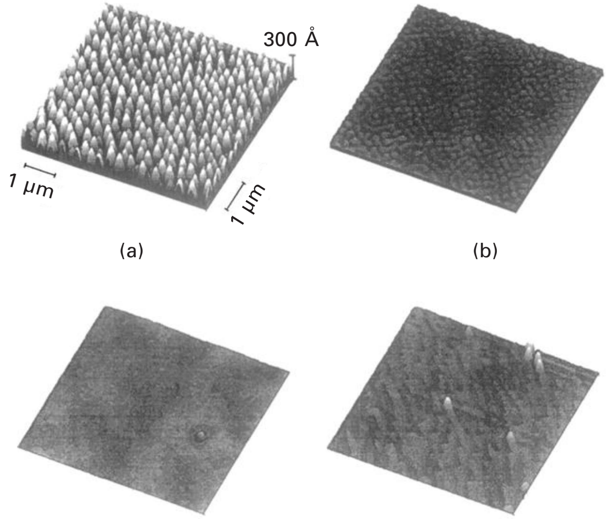
Based upon the discovery of low dislocation nucleation rates in low-mismatch SiGe/Si(001) at 900°C coupled with the high dislocation velocities expected for such a high temperature, Fitzgerald *et al.* (1991) reasoned that a low number of threads could relax a large amount of strain at high temperature. Note that the degree of relaxation depends on the total misfit length and not explicitly on threading dislocation density, TDD (Fitzgerald, 1991). In principle, another low mismatch layer could then be grown, and another, leading to a graded structure where the cap is actually highly mismatched from the substrate. The realization of relaxed SiGe buffers with greatly reduced TDD (Fig. 13.3) spurred interest in strained Si FETs (Fitzgerald *et al.*, 1991; Tuppen *et al.*, 1991). Through the use of compositional grading at high temperatures, Fitzgerald *et al.* (1991) at AT&T lowered the TDD in the active area of a strained Si modulation-doped structure from 10^8 cm^{-2} to 10^6 cm^{-2}, allowing the record for 4 K electron mobility to be improved from 19,000 to 96,000 cm^2/Vs. Within a year, the AT&T group increased their record mobilities to 170,000 cm^2/Vs (Fitzgerald *et al.*, 1992). Subsequently, a group at Daimler-Benz employed the graded buffer technique and also attained electron mobilities of 170,000 cm^2/Vs (Schäffler *et al.*, 1992). Many other concepts for attaining relaxed SiGe buffers on Si wafers, such as compliant substrate and low-temperature buffer techniques, have been proposed and demonstrated over the years. However, compositional grading remains the most established technique for attaining low-defect density, fully relaxed SiGe alloys on a Si substrate, as demonstrated by the thorough physical understanding that exists in the literature and the number of academic and commercial entities that have successfully demonstrated SiGe graded buffers.



13.3 Cross-sectional TEM of SiGe/Si(001) compositionally graded buffer showing low-TDD, highly relaxed cap layer.

13.2.5 Effect of sign of strain: compressive vs tensile

Although the vast majority of work on SiGe has placed the epilayer in compressive strain, the effect of reversing the sign of strain on dislocations in SiGe can be dramatic. Wegscheider *et al.* (1990) pointed out that relaxation of tensile SiGe on Ge(001) films took place via nucleation and glide of 90° partial dislocations as opposed to 60° dislocations, resulting in films with significant densities of stacking faults and microtwins. Wegscheider's discovery foreshadowed the observation of high densities of stacking faults and microtwins in strained Si layers. Another startling discovery concerning the sign of strain came when Xie *et al.* (1994) reported that while Ge on $Si_{0.5}Ge_{0.5}$ spontaneously self-assembles into nanostructures, Si on $Si_{0.5}Ge_{0.5}$ does not, instead remaining flat and eventually developing stacking faults (Fig. 13.4). The dependence of morphology on the sign of strain cannot be predicted by continuum mechanics approaches, where details of the surface such as reconstruction and steps are ignored. Xie *et al.* (1994) used molecular dynamics simulations to show that compressive strain caused a reduction in surface step energy, while tensile strain raised the surface step energy. The increase in step energy would then discourage the Stranski–Krastanov growth mode, consistent with experimental observation.



13.4 AFM images of $Si_{0.5}Ge_{0.5}$ under (a) 2% compressive strain, (b) 1% compressive strain, (c) 0.5% compressive strain, and (d) 2% tensile strain. Reprinted with permission from Xie *et al.*, *Physical Review Letters*, vol. 73, 3006, 1994. Copyright 1994 by the American Physical Society.

13.2.6 Brief summary of historical observations

The high nucleation barrier for half-loop nucleation combined with the low concentration of fixed dislocation sources dominated many of the early observations of extended defects in compressive SiGe materials on Si(001). Recapping, People and Bean were able to observe dislocation-free SiGe films, to within the limitations of their observation techniques, grown several times the thermodynamically predicted critical thickness (People and Bean, 1986; Bean *et al.*, 1984). Through the use of substrate patterning, Fitzgerald *et al.* (1990) pointed out that the metastability can persist at very high growth temperatures of 900°C provided a fixed heterogeneous source is not present. The same metastability allowed Eaglesham and Cerullo (1990) to observe dislocation-free Ge islands grown on Si(001) via the Stranski–Krastanov mechanism. At their peak, these islands can be many times thicker than the predicted critical thickness. Recognizing that large strain relaxations can be accomplished with a low density of threads and long misfit segments by growth at high temperature, both Fitzgerald and Tuppen demonstrated SiGe graded buffers with low TDD (Fitzgerald *et al.*, 1991; Tuppen *et al.*, 1991). Although the majority of research on extended defects in SiGe and SiGe nanostructures had been on Si(001), several important changes were noted when either the sign of strain was reversed or the surface orientation was changed. First, SiGe films on (110) and (111) Si were prone to develop stacking faults due to the nucleation and glide of 90° Shockley partial dislocations unlike their counterparts grown on Si(001) (Hull *et al.*, 1991b). Similarly, tensile SiGe films grown on Ge (001) were also found to relax by glide of Shockley partials (Wegscheider *et al.*, 1990). Finally, while Ge on $Si_{0.5}Ge_{0.5}$ was found to exhibit S–K growth, Si on $Si_{0.5}Ge_{0.5}$ was found to remain flat and eventually relax by glide of 90° Shockley partials (Xie *et al.*, 1994).

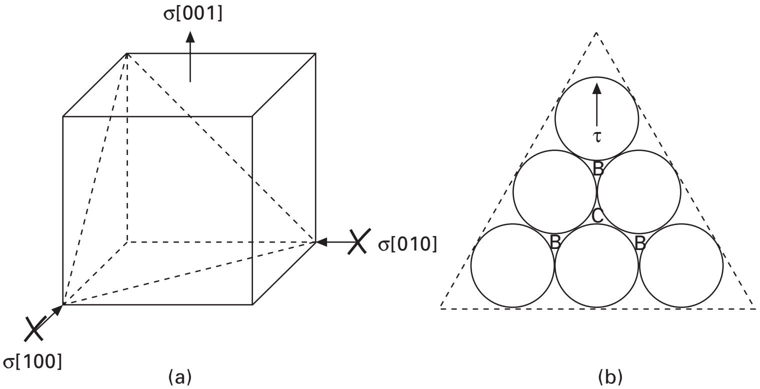
13.3 Application of the Thompson tetrahedron to extended defects in silicon–germanium (SiGe)

Unlike the compressive case, highly tensile growth on (001) substrates results in a planar, dislocated wetting layer with stacking faults (Pachinger *et al.*, 2007; Xie *et al.*, 1994). As a result, it remains impossible to grow dislocation-free tensile Si SAQDs on Ge(001) surfaces (Pachinger *et al.*, 2007). One fortuitous and oft-overlooked aspect of compressive growth on (001) is that stacking fault formation in biaxially compressed films is highly unlikely, as can be predicted using the Thompson tetrahedron (Thompson, 1953). Most dislocations in diamond cubic and zinc blende semiconductors are 60° $a/2[101]$ total dislocations that are slightly dissociated into 90° and 30° Shockley partial dislocations bound by a narrow stacking fault 'ribbon'

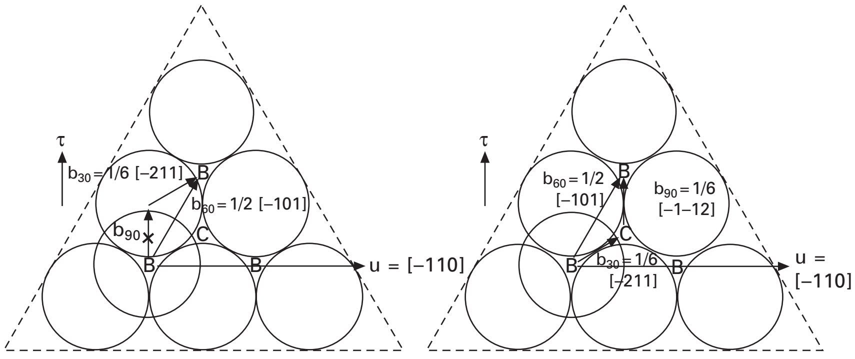
(Fig. 13.2): $a/2[101] = a/6[211] + a/6[112]$. The passage of a partial dislocation creates a stacking fault by definition, since the Burgers vector does not correspond to a lattice translation vector. Therefore, for a dissociated total dislocation, there is an explicitly defined sequence in which the partials must pass (i.e. one leads and one trails) in order to maintain close packing and to avoid creation of a high-energy A|A stacking fault.

Let us first schematically illustrate the conversion of a biaxial *compression* ($\sigma[100] = \sigma[010]$, $\sigma[001] = 0$) into a uniaxial *tension* along [001] ($\sigma[100] = \sigma[010] = 0$, $\sigma[001] > 0$) by addition of a hydrostatic stress tensor (hydrostatic stresses produce no shear strain and thus have no effect on dislocation motion) whose terms are equal and opposite to the original biaxial compression (Fig. 13.5(a)) (Kvam and Hull, 1993). The uniaxial tension can then be resolved into a shear stress τ on a {111} glide plane as shown in Fig. 13.5(b) where gray circles denote atoms in the A plane, and the B and C sites are labeled. For 60° $a/2[101]$ total dislocations in compressive layers on (001), the 90° partial cannot glide first because doing so would create a high-energy A|A stacking fault as in Fig. 13.6(a). The 30° partial must lead, but the 90° partial experiences a larger shear stress (i.e. its Burgers vector and the shear stress vector τ are aligned in parallel, Fig. 13.6(b)) and thus always trails the 30° partial closely, preventing wide stacking faults from ever forming (Marée *et al.*, 1987).

In contrast, for biaxial tension, the sign of τ is reversed (Fig. 13.7) causing the 90° Shockley partial to lead and to experience greater resolved shear stress than both the 30° Shockley partial and the 60° total dislocation. Thus a 60° total dislocation in [001]-oriented biaxial tension can become widely dissociated with a large stacking fault between the two partials, as is

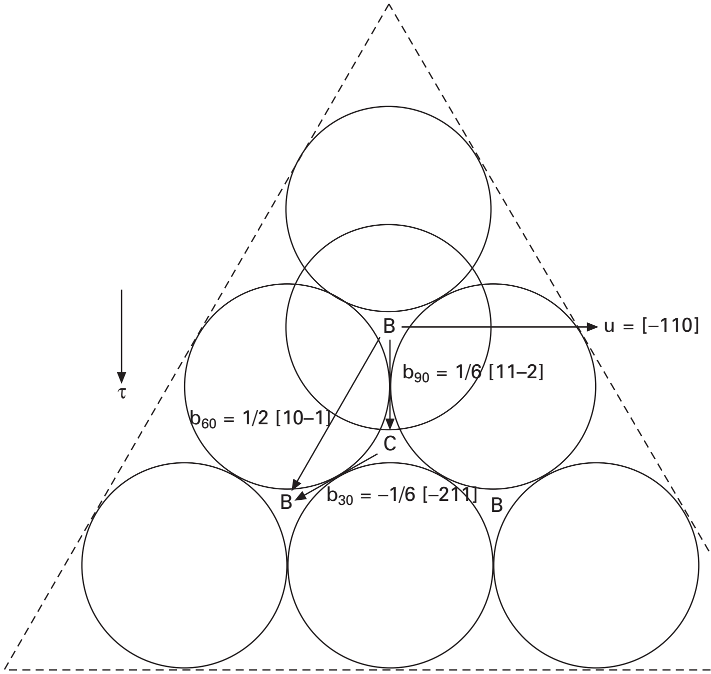


13.5 (a) Conversion of a biaxial compressive stress state ($\sigma[100] = \sigma[010]$, $\sigma[001] = 0$) into a uniaxial tensile stress ($\sigma[100] = \sigma[010] = 0$, $\sigma[001] > 0$) by addition of a hydrostatic stress tensor; (b) resolution of the uniaxial tensile stress $\sigma[001]$ into a shear stress τ on {111}.


τ
b_{30} = 1/6 [–211]
B
b_{60} = 1/2 [–101]
b_{90}
C
B
B
u = [–110]
τ
b_{60} = 1/2
[–101]
B
b_{90} = 1/6
[–1–12]
C
B
b_{30} = 1/6
[–211]
B
u =
[–110]

(a) (b)

13.6 Shear strain on a {111} glide plane resulting from (001)-oriented biaxial compression. (a) Creation of a high-energy A|A stacking fault would occur if the 90° Shockley partial were to lead and is therefore highly unlikely under biaxial compression. (b) When the 30° partial leads, a lower-energy A|C stacking fault is created, but the fault never widens since the 90° Shockley partial experiences greater shear stress and is always pushed against the 30° Shockley partial.


B
u = [–110]
τ
b_{90} = 1/6 [11–2]
b_{60} = 1/2 [10–1]
C
B
b_{30} = –1/6 [–211]
B

13.7 Biaxial tension causes the 90° Shockley partial to lead and to experience greater resolved shear stress than the 30° partial, lowering the barrier to dislocation nucleation and allowing large stacking faults to form. The lowered barrier also inhibits the ability to form dislocation-free QDs.

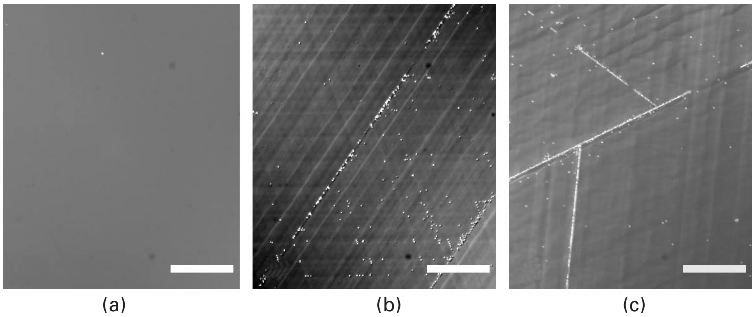
commonly seen in strained Si grown on SiGe(001) (Kimura *et al.*, 2006; Xie *et al.*, 1994; Wegscheider *et al.*, 1990). A microtwin can then be formed by the shearing action of numerous partial dislocations on adjacent glide planes (Wagner, 1998; Wegscheider *et al.*, 1990). Furthermore, if we normalize to the Burgers vector magnitude, a 90° Shockley partial has lower strain energy and a larger strain-relieving component than a 60° total dislocation (higher Schmid factor), since the 90° partial resolves only into tilt and strain-relieving components; the 60° total dislocation possesses an additional screw component. The higher Schmid factor also means that the nucleation barrier of a 90° partial should be significantly lower than that of a 60° total dislocation (Marée *et al.,* 1987). Therefore, the large barrier to dislocation nucleation, which aids the formation of dislocation-free compressive QDs on (001), is not present for biaxial tension. The observation of a thick, dislocated wetting layer with stacking faults in tensile Si on Ge(001) can thus be correctly predicted using the Thompson tetrahedron.

For compressive layers on (111) and (110) substrates, the 90° partial dislocation leads the 30° partial and also experiences a larger resolved shear stress (Kohler *et al.,* 1991; Hull *et al.,* 1991b). Therefore, unlike the familiar case of compressive films on (001), where dislocations tend to remain narrowly dissociated, 60° dislocations in biaxially compressive films grown on (111) and (110) can become widely dissociated, resulting in high densities of stacking faults. These stacking faults can have significant consequences on the global relaxation and TDD of such films, as discussed in Section 13.4 below.

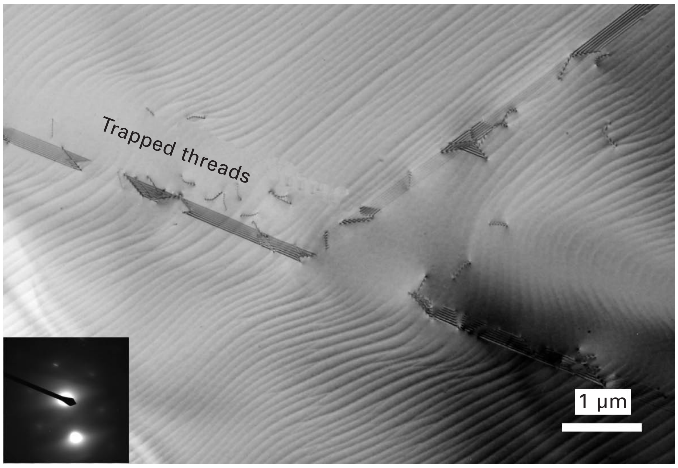
13.4 Current topics

13.4.1 Feasibility of attaining high-quality relaxed SiGe on Si(111) and Si(110)

The ability to grow both strained and relaxed SiGe films of high structural perfection on arbitrary substrate orientations would allow the coupled effects of surface orientation and strain on carrier mobility to be studied in much greater detail. However, Fig. 13.8 shows Nomarski micrographs of Schimmel-etched surfaces for low-mismatch SiGe grown on Si(001)/(110)/(111), and reveals that the TDD on (001) (Fig. 13.8(a)) is more than an order of magnitude lower than the other cases. In addition to numerous isolated pits, the (111) and (110) films (Figs 13.8(b)) and (c)) show bright lines of closely spaced pits that appear to trap threads in their vicinity. Figure 13.9 shows a planar-view transmission electron micrograph of one such pileup in $Si_{0.97}Ge_{0.03}$ on Si(111), where it can be seen that numerous threads are arrested near planar defects with stacking fault-like contrast. Since planar defects apparently block the glide of threads, further nucleation events are


(a) (b) (c)

13.8 Nomarski micrographs of Schimmel-etched $Si_{0.97}Ge_{0.03}$ on (a) Si(001), (b) Si(110), (c) Si(111), and (d) Si(112). Note the propensity for threads to become trapped near the bright line features (microtwins) in (b) and (c). Scale bars represent 25 μm. Adapted from Lee *et al*. (2006).



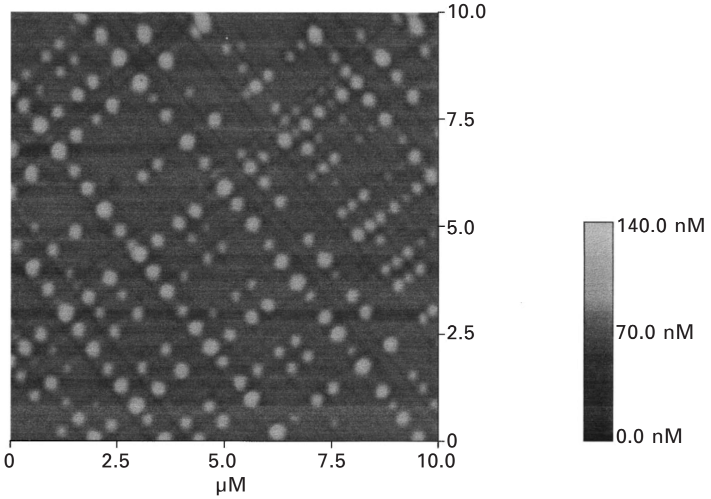
13.9 PVTEM of $Si_{0.97}Ge_{0.03}$ on Si(111) (same growth as in Fig. 13.8(c)) showing high numbers of threads piled up near planar defects. The wavy line contrast is caused by surface step bunches. Adapted from Lee *et al*. (2006).

needed to reduce the elastic energy in the film, and the net result is a global increase in TDD.

The primary problem with compressive films grown on (111) and (110) is that the nucleation and glide of 90° Shockley partials is a significant mechanism for strain relief. For these orientations, the glide planes are oriented such that the 90° partial leads and also has a larger Schmid factor than the 60° total dislocation (0.31 vs. 0.27 for (111) substrates and 0.47 vs. 0.41 for (110) substrates). Since microtwin defects in SiGe grown on Si(110) and (111) appear to arrest the glide of other dislocations, the nucleation and propagation of planar defects must be avoided in order to attain high-quality epitaxial materials (Lee *et al.,* 2006). While grading rates of 10–20% Ge/µm can be successfully used for the growth of low-TDD (001) buffers, it is clear that the (111) and (110) substrates cannot tolerate such high strain rates (Lee *et al.,* 2006; Destefanis *et al.,* 2009). Still, it remains possible that high-quality films could be attained on (111) or (110) wafers by using extremely slow grading rates (e.g. 1% Ge/µm), though such thick buffers may be technologically unfeasible.

13.4.2 Use of misfit dislocations as a template for quantum dot growth

The long-range strain field of misfit dislocations is known to be responsible for the characteristic cross-hatch surface roughness pattern of relaxed SiGe on Si (Fitzgerald *et al.,* 1992). In most work, the cross-hatch roughness is at best a nuisance to be dealt with via ex situ polishing before subsequent device processing (Currie *et al.,* 2001). At worst, the extreme height fluctuations can block dislocation glide and lead to threading dislocation pileups (Currie *et al.,* 1998). Xie *et al.* (1997) and Shiryaev *et al.* (1997) made the nearly simultaneous discovery that the strain fields from misfit dislocations could be used to guide growth of Ge QDs (Fig. 13.10). The dots were found to nucleate in rectangular arrays resembling the array of misfit dislocations underneath



13.10 10 µm × 10 µm AFM image of the sample with 1.0 nm of Ge coverage grown at 750°C showing 2D array formation. From Xie *et al.* (1997). Copyright 1997 by the American Institute of Physics.

due to modulations in surface chemical potential. Elaborating further, Kim *et al*. (2003) found that three types of surface sites could be identified based upon the tendency of Ge dots to nucleate on them; the intersection between two misfit dislocations was found to be the most preferred QD nucleation site (or the lowest chemical energy site for adatoms), the area above a single misfit dislocation was next, and the field between misfit dislocations was the least preferred.

13.4.3 SiGe/Si quantum dot superlattice solar cell

One area of interest for SiGe/Si QDs is in the quest to attain thin, high-efficiency Si solar cells in order to lower the price per peak watt of photovoltaic energy. For example, Brendel and Scholten (1999) have computed that a Si solar cell only 1–3 μm thick could attain 17–18% efficiency if nearly perfect light trapping could be attained. Such thin Si cells can enable a significant reduction in cost per watt, because the reduced base thickness relaxes the need for a long minority-carrier lifetime and makes them less sensitive to defects related to process-induced damage and contamination. Besides lower cost, a thin Si cell might possess additional advantages over its bulk counterpart, such as superior radiation resistance, lower launch weight for space applications, and mechanical flexibility. However, thin crystalline Si cells suffer from low short-circuit current density (J_{sc}) and low efficiency, because the absorption coefficient of Si is far less than that of direct-gap semiconductors for much of the solar spectrum. While light management techniques such as surface texturing (Verlinden *et al.,* 1992), back-surface reflectors (Zeng *et al.,* 2006), and surface plasmon engineering (Pillai *et al.,* 2007) can improve absorption, it remains difficult for thin c-Si cells to attain high J_{sc}.

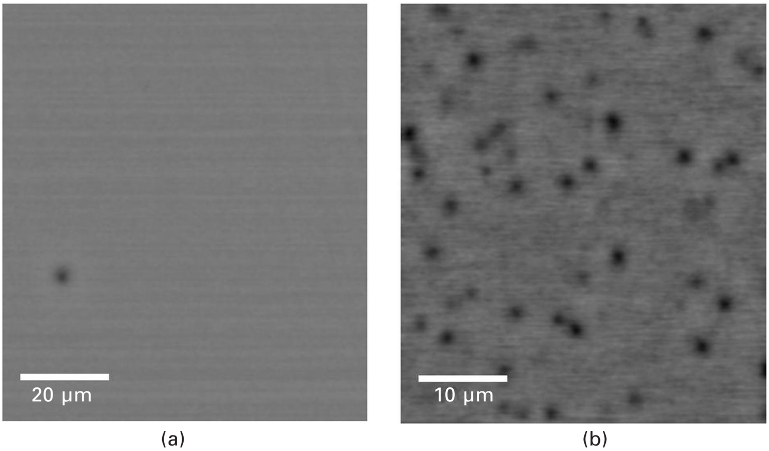
One way to improve near-infrared (NIR) absorption in thin Si solar cells (E_g = 1.12 eV) is to incorporate lower-bandgap Ge (E_g = 0.67 eV) or SiGe (0.67 eV $\leq E_g \leq$ 1.12 eV) into the base region. Ge and SiGe can spontaneously self-assemble into dislocation-free quantum dots (QDs) when grown on Si(001), allowing integration of IR-sensitive nanostructures into Si optoelectronic devices. To increase optical absorption and sensitivity, the QDs can be capped and planarized with Si, allowing a subsequent layer of QDs to be deposited (Abstreiter *et al.,* 1996). Repetition of these steps leads to the formation of quantum dot superlattices (QDSLs) where multiple layers of QDs are separated from one another by thin Si barriers (Mateeva *et al.,* 1997).

Although SiGe dots can be grown dislocation-free well beyond their critical thickness, eventually there comes a point where dislocation nucleation begins. Lee *et al*. (2010) compared 40-period SiGe/Si QDSL solar cells with two different QD densities: 5.5×10^9 cm^{-2} and 8.5×10^9 cm^{-2}. X-ray scans about (004) revealed that the cell with denser QDs had a 29% higher splitting of the zeroth-order superlattice peak from the substrate peak than the cell with sparser dots. Since the superlattice periods of both QDSLs were nearly identical (44 ± 2 nm, calculated from superlattice fringe spacings), it was deduced that the cell with denser QDs possessed ~29% higher Ge content on average (Baribeau and Houghton, 1991) than the other cell.

In order to analyze the photovoltaic performance of the QDSL and to simulate a solar cell where the QDSL has been grown on low-lifetime Si or transferred to a foreign host substrate, Lee *et al.* (2010) deposited their SiGe/Si QDSLs onto Czochralski-grown p^+ Si wafers ($N_A \sim 10^{20}$ cm^{-3}). Since the minority carrier diffusion length in such heavily doped Si is ≤1 μm (Tyagi and Van Overstraeten, 1983), very few of the electron-hole pairs generated within the bulk are collected as photocurrent. Besides serving as a contact and a natural back-surface field, the p^+ Si substrate can be considered to be optically and electrically inactive, ensuring that nearly all the photocurrent was generated in the epitaxial layers.

The motivation for growing QDSLs with higher QD density was to boost infrared photocurrent without dramatically increasing the thickness, but the two cells were found to possess very similar values of J_{sc} and EQE. In addition, the cell with denser QDs suffered a 120 mV drop in V_{oc} from the V_{oc} of 0.51 V observed in the cell with sparser QDs. Some reduction in V_{oc} for the cell with denser dots is expected, since the higher density of QDs and higher overall Ge content should, on average, lead to a lower bandgap. However, a lower bandgap should be accompanied by an increase in J_{sc}, and Lee *et al.* (2010) therefore performed EBIC analysis on the two cells to compare the density of electrically active defects. It was found that while the cell with a QD density of 5.5×10^9 cm^{-2} had a dark spot density (DSD) of ~10^4 cm^{-2}, the cell with a QD density of 8.5×10^9 cm^{-2} had a DSD of ~3×10^6 cm^{-2} (Fig. 13.11; note that the magnification of Fig. 13.11(b) is twice that of Fig. 13.11(a)). The dark spots in these EBIC images likely arise from recombination at threading dislocations (Leamy, 1982), and DSD should correlate well with TDD, provided that the spots are spaced sufficiently to be individually resolved. Vernon *et al.* (1990) found that in otherwise identical GaAs cells, V_{oc} and J_{sc} could be degraded by 20% and 10%, respectively, as TDD rose from 5×10^5 to 7×10^7 cm^{-2}. Thus the low V_{oc} and relatively unchanged J_{sc} of the cell with denser QDs can likely be attributed to a combination of its reduced bandgap and higher TDD. The differences in TDD between the two cells indicate that the dislocation nucleation rate in SiGe/Si QDSLs depends sensitively on QD density, and that a relatively low density must be maintained in order to avoid dislocation-related V_{oc} reduction.

For QDSL cells with a relatively low QD density, it is possible that continuing to increase the base thickness could lead to higher responsivity and



(a) (b)

13.11 10 kV EBIC images of (a) 40-period QDSL, QD density 5.5×10^9 cm^{-2} and (b) 40-period QDSL, QD density 8.5×10^9 cm^{-2}. Note that the magnification of (b) is twice that of (a). The dark spot density was ~10^4 cm^{-2} in (a) and 3×10^6 cm^{-2} in (b). From Lee *et al.* (2010), copyright 2010 by Elsevier.

higher J_{sc}, paving the way towards thin, high-efficiency Si-based photovoltaic cells. Future work in this area should focus on QDSLs with more than 100 periods that are deposited on strain-relaxed $Si_{1-x}Ge_x$ buffers in order to avoid dislocation-related V_{oc} reduction.

13.5 Future trends

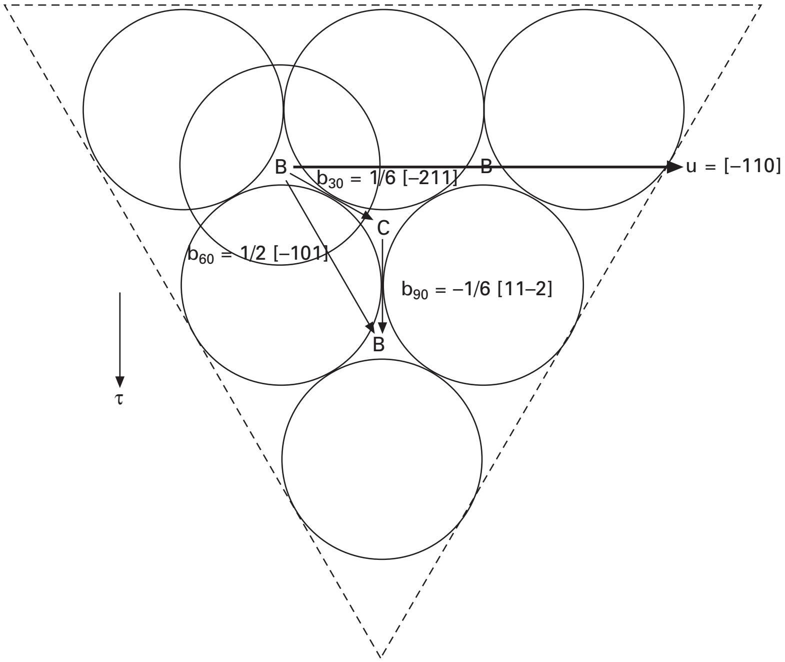
The topic of extended defects in SiGe nanostructures will continue to be rich in the foreseeable future, with several interesting directions presenting themselves. Here we discuss issues and possible opportunities for the study of defects in quantum dots and nanowires.

13.5.1 Defects in SiGe heterostructure nanowires

Recently, SiGe nanowires have been proposed as future functional elements for nanoscale electronics and sensing (Xiang *et al.,* 2006). Despite the widely held belief that such nanostructures are more tolerant to mismatch strain due to the possibility of elastic strain sharing (Wu *et al.,* 2002), several recent reports have highlighted the presence of dislocations in both Ge/Si core/shell nanowire heterostructures (Goldthorpe *et al.,* 2008) and Si/Ge axial heterostructures (Shimizu *et al.,* 2009). In addition, Si nanowires are known to exhibit both wurtzite and diamond cubic polytypes due to grown-in stacking faults (Lopez *et al.,* 2009). The ability to predict when dislocations and stacking faults should and should not form is still needed, as is deeper understanding of the consequences of these defects on device operation.

13.5.2 Self-assembled Si quantum dots

Self-assembled Si QDs may be of future interest for single-electron transistors for quantum information processing (Pachinger *et al.,* 2007). As mentioned before, earlier attempts to grow Si QDs on Ge(001) resulted in dislocated layers and dots (Pachinger *et al.,* 2007). However, when biaxial tension is applied to (110) or (111) growth, instead of changing the sign of τ, we rotate the Thompson tetrahedron itself (Fig. 13.12) and find that the 30° partial should lead while the trailing 90° partial experiences greater shear stress on both (110) and (111) (Kvam and Hull, 1993). From the standpoint of which partial leads and which partial trails, biaxial tension on (110) and (111) is identical to the case of biaxial compression on (001). Thus it may be predicted that dislocation-free QDs can be grown in tension on (110) and (111) substrates. To the author's knowledge the only existing experimental

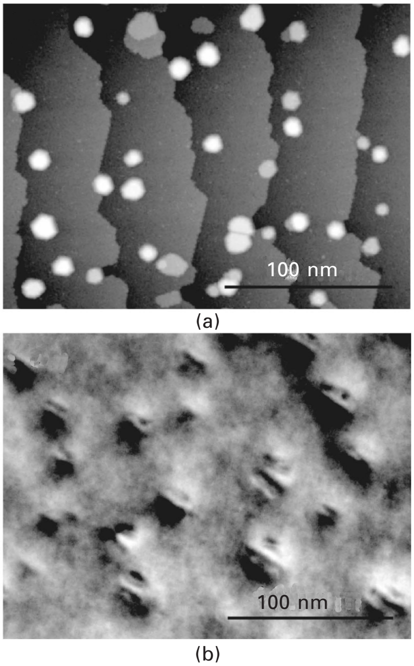


13.12 Use of the Thompson tetrahedron aids visualization of the shear stresses on (111) glide planes resulting from biaxial tension on (110)- and (111)-oriented substrates. The magnitude of resolved shear differs for (110) and (111) due to the different Schmid factors, but the 30° Shockley partial leads for both surface orientations.

data suggesting the validity of this prediction is the work of Raviswaran *et al*. (2001) who grew dislocation-free Si QDs on Ge(111) (Figure 13.13). Interestingly, the Si/Ge(111) QDs exhibited many differences from the more familiar case of Ge/Si(001), such as the lack of a wetting layer and the ability to attain extremely high dot densities of 5×10^{11} cm^{-2}. According to the analysis presented in this chapter, it should also be possible to grow dislocation-free Si dots on Ge(110), as well.

13.6 Acknowledgments

The author is grateful to Ya-Hong Xie and David Cahill for permission to reprint figures from their papers. The author also thanks Yuncheng Song for assistance with graphics.



13.13 (a) STM image of coherent islands formed on Ge(111) by the deposition of two equivalent bilayers (BLs) of Si at 500°C. (b) Dark-field strain-contrast plan-view TEM images (g = ⟨220⟩) of islands formed by deposition of 2.8 BL Si on Ge(111) at 550°C. About half of the islands are coherent and half are dislocated. Reprinted with permission from Raviswaran *et al*., *Physical Review B*, vol. 63, 125314, 2001. Copyright 2001 by the American Physical Society.

13.7 References

Abstreiter, G., Schittenhelm, P., Engel, C., Silveira, E., Zrenner, A., Meertens, D., Jager, W. (1996), Growth and characterization of self-assembled Ge-rich islands on Si. *Semiconductor Science and Technology*, **11**, 1521–1528.

Baribeau, J.-M., Houghton, D. C. (1991), Application of X-ray diffraction techniques to the structural study of silicon based heterostructures. *Journal of Vacuum Science and Technology B: Microelectronics and Nanometer Structures*, **9**, 2054–2058.

Bean, J. C., Feldman, L. C., Fiory, A. T., Nakahara, S., Robinson, I. K. (1984), Ge_xSi_{1-x}/Si strained-layer superlattice grown by molecular beam epitaxy. *Journal of Vacuum Science and Technology A: Vacuum, Surfaces, and Films*, **2**, 436–440.

Brendel, R., Scholten, D. (1999), Modeling light trapping and electronic transport of waffle-shaped crystalline thin-film Si solar cells. *Applied Physics A: Materials Science* and *Processing*, **69**, 201–213.

Currie, M. T., Samavedam, S. B., Langdo, T. A., Leitz, C. W., Fitzgerald, E. A. (1998), Controlling threading dislocation densities in Ge on Si using graded SiGe layers and chemical–mechanical polishing. *Applied Physics Letters*, **72**, 1718–1720.

Currie, M. T., Leitz, C. W., Langdo, T. A., Taraschi, G., Fitzgerald, E. A., Antoniadis, D. A. (2001), Carrier mobilities and process stability of strained Si *n*- and *p*-MOSFETs on SiGe virtual substrates. *Journal of Vacuum Science and Technology B: Microelectronics and Nanometer Structures*, **19**, 2268–2279.

Destefanis, V., Hartmann, J.M., Abbadie, A., Papon, A.M., Billon, T. (2009), Growth and structural properties of SiGe virtual substrates on Si(1 0 0), (1 1 0) and (1 1 1). *Journal of Crystal Growth*, **311**, 1070–1079.

Duesbery, M. S., Richardson, G. Y. (1991), The dislocation core in crystalline materials. *Critical Reviews in Solid State and Materials Sciences*, **17**, 1–46.

Eaglesham, D. J., Cerullo, M. (1990), Dislocation-free Stranski–Krastanow growth of Ge on Si(100). *Physical Review Letters*, **64**, 1943–1946.

Fiorenza, J. G., Braithwaite, G., Leitz, C. W., Currie, M. T., Yap, J., Singaporewala, F., Yang, V. K., Langdo, T. A., Carlin, J., Somerville, M., Lochtefeld, A., Badawi, H., Bulsara, M. T. (2004), Film thickness constraints for manufacturable strained silicon CMOS. *Semiconductor Science and Technology*, **19**, L4–L8.

Fitzgerald, E. A. (1991), Dislocations in strained-layer epitaxy: theory, experiment, and applications. *Materials Science Reports*, **7**, 87–140.

Fitzgerald, E. A., Kirchner, P. D., Proano, R., Pettit, G. D., Woodall, J. M., Ast, D. G. (1988), Elimination of interface defects in mismatched epilayers by a reduction in growth area. *Applied Physics Letters*, **52**, 1496–1498.

Fitzgerald, E. A., Xie, Y. H., Brasen, D., Green, M. L., Michel, J., Freeland, P. E., Weir, B. E. (1990), Elimination of dislocations in heteroepitaxial MBE and RTCVD Ge_xSi_{1-x} grown on patterned Si substrates. *Journal of Electronic Materials*, **19**, 949–955.

Fitzgerald, E. A., Xie, Y. H., Green, M. L., Brasen, D., Kortan, A. R., Mii, Y. J., Michel, J., Weir, B. E., Feldman, L. C., Kuo, J. M. (1991), Strain-free Ge_xSi_{1-x} with low threading dislocation densities grown on Si substrates. In Bean, J. C., Iyer, S. S. and Wang, K. L. (eds), *Mater. Res. Soc. Symp. Proc.*, **220**, 211–215.

Fitzgerald, E. A., Xie, Y. H., Monroe, D., Silverman, P. J., Kuo, J. M., Kortan, A. R., Thiel, F. A., Weir, B. E. (1992), Relaxed Ge_xS_{1-x} structures for III–V integration with Si and high mobility two-dimensional electron gases in Si. *Journal of Vacuum Science and Technology B: Microelectronics and Nanometer Structures*, **10**, 1807–1819.

Goldthorpe, I. A., Marshall, A. F., McIntyre, P. C. (2008), Synthesis and strain relaxation of Ge-core/Si-shell nanowire arrays. *Nano Letters*, **8**, 4081–4086.

Green, M. L., Weir, B. E., Brasen, D., Hsieh, Y. F., Higashi, G., Feygenson, A., Feldman, L. C., Headrick, R. L. (1991), Mechanically and thermally stable Si–Ge films and heterojunction bipolar transistors grown by rapid thermal chemical vapor deposition at 900°C. *Journal of Applied Physics*, **69**, 745–751.

Grillot, P. N., Ringel, S. A., Michel, J., Fitzgerald, E. A. (1996), Structural, electronic, and luminescence investigation of strain-relaxation induced electrical conductivity type conversion in GeSi/Si heterostructures. *Journal of Applied Physics*, **80**, 2823–2832.

Higgs, V., Kittler, M. (1993), Investigation of the recombination activity of misfit dislocations in Si/SiGe epilayers by cathodoluminescence imaging and the electron beam induced current technique. *Applied Physics Letters*, **63**, 2085–2087.

Higgs, V., Lightowlers, E. C., Tajbakhsh, S., Wright, P. J. (1992), Cathodoluminescence imaging and spectroscopy of dislocations in Si and $Si_{1-x}Ge_x$ alloys. *Applied Physics Letters*, **61**, 1087–1089.

Hull, R., Bean, J. C. (1992), Misfit dislocations in lattice-mismatched epitaxial films. *Critical Reviews in Solid State and Materials Sciences*, **17**, 507–546.

Hull, R., Bean, J. C., Peticolas, L., Bahnck, D. (1991a), Growth of Ge_xSi_{1-x} alloys on Si(110) surfaces. *Applied Physics Letters*, **59**, 964–966.

Hull, R., Bean, J. C., Peticolas, L., Xie, Y. H., Hsieh, Y. F. (1991b), Growth of Ge_xSi_{1-x}/Si alloys on Si (100), (110) and (111) surfaces. In Bean, J. C., Iyer, S. S., and Wang, K. L. (eds), *Mater. Res. Soc. Symp. Proc.*, **220**, 153–159.

Kamins, T. I., Carr, E. C., Williams, R. S., Rosner, S. J. (1997), Deposition of three-dimensional Ge islands on Si(001) by chemical vapor deposition at atmospheric and reduced pressures. *Journal of Applied Physics*, **81**, 211–219.

Kim, H. J., Zhao, Z. M., Xie, Y. H. (2003), Three-stage nucleation and growth of Ge self-assembled quantum dots grown on partially relaxed SiGe buffer layers. *Physical Review B*, **68**, 205312+.

Kimura, Y., Sugii, N., Kimura, S., Inui, K., Hirasawa, W. (2006), Generation of misfit dislocations and stacking faults in supercritical thickness strained-Si/SiGe heterostructures. *Applied Physics Letters*, **88**, 031912+.

Kohler, U., Jusko, O., Pietsch, G., Muller, B., Henzler, M. (1991), Strained-layer growth and islanding of germanium on Si(111)-(7 × 7) studied with STM. *Surface Science*, **248**, 321–331.

Kvam, E. P., Hull, R. (1993), Surface orientation and stacking fault generation in strained epitaxial growth. *Journal of Applied Physics*, **73**, 7407–7411.

Leamy, H. J. (1982), Charge collection scanning electron microscopy. *Journal of Applied Physics*, **53**, R51–R80.

Lee, M. L. (2005), *unpublished*.

Lee, M. L., Fitzgerald, E. A. (2004), Electron mobility characteristics of n-channel metal-oxide-semiconductor field-effect transistors fabricated on Ge-rich single- and dual-channel SiGe heterostructures. *Journal of Applied Physics*, **95**, 1550–1555.

Lee, M. L., Antoniadis, D., Fitzgerald, E. A. (2006), Challenges in epitaxial growth of SiGe buffers on Si (111), (110), and (112). *Thin Solid Films*, **508**, 136–139.

Lee, M. L., Dezsi, G., Venkatasubramanian, R. (2010), Analysis of SiGe/Si quantum dot superlattices grown by low-pressure chemical vapor deposition for thin solar cells. *Thin Solid Films*, **518**, S76–S79.

LeGoues, F. K., Copel, M., Tromp, R. M. (1990), Microstructure and strain relief of Ge films grown layer by layer on Si(001). *Physical Review B*, **42**, 11690–11700.

Lopez, F. J., Hemesath, E. R., Lauhon, L. J. (2009), Ordered stacking fault arrays in silicon nanowires. *Nano Letters*, **9**, 2774–2779.

Low, T., Li, M. F., Shen, C., Yeo, Y. C., Hou, Y. T., Zhu, C., Chin, A., Kwong, D. L. (2004), Electron mobility in Ge and strained-Si channel ultrathin-body metal-oxide semi conductor field-effect transistors. *Applied Physics Letters*, **85**, 2402–2404.

Luan, H.-C., Lim, D. R., Lee, K. K., Chen, K. M., Sandland, J. G., Wada, K., Kimerling, L. C. (1999), High-quality Ge epilayers on Si with low threading-dislocation densities. *Applied Physics Letters*, **75**, 2909–2911.

Marée, P. M. J., Barbour, J. C., van der Veen, J. F., Kavanagh, K. L., Lieuwma, C. W. T. Bulle, Viegers, M. P. A. (1987), Generation of misfit dislocations in semiconductors. *Journal of Applied Physics*, **62**, 4413–4420.

Marshall, A. F., Aubertine, D. B., Nix, W. D., McIntyre, P. C. (2005), Misfit dislocation dissociation and Lomer formation in low mismatch SiGe/Si heterostructures. *Journal of Materials Research*, **20**, 447–455.

Mateeva, E., Sutter, P., Bean, J. C., Lagally, M. G. (1997), Mechanism of organization of three-dimensional islands in SiGe/Si multilayers. *Applied Physics Letters*, **71**, 3233–3235.

Matthews, J. W., Blakeslee, A. E. (1974), Defects in epitaxial multilayers: I. Misfit dislocations. *Journal of Crystal Growth*, **27**, 118–125.

Pachinger, D., Groiss, H., Lichtenberger, H., Stangl, J., Hesser, G., Schäffler, F. (2007), Stranski–Krastanow growth of tensile strained Si islands on Ge (001). *Applied Physics Letters*, **91**, 233106.

People, R., Bean, J. C. (1985), Calculation of critical layer thickness versus lattice mismatch for Ge_xSi_{1-x}/Si strained-layer heterostructures. *Applied Physics Letters*, **47**, 322–324.

People, R., Bean, J. C. (1986), Erratum: Calculation of critical layer thickness versus lattice mismatch for Ge_xSi_{1-x}/Si strained-layer heterostructures [*Appl. Phys. Lett.*, **47**, 322 (1985)]. *Applied Physics Letters*, **49**, 229.

Pillai, S., Catchpole, K. R., Trupke, T., Green, M. A. (2007), Surface plasmon enhanced silicon solar cells. *Journal of Applied Physics*, **101**, 093105.

Radzimski, Z. J., Zhou, T. Q., Buczkowski, A., Rozgonyi, G. A., Finn, D., Hellwig, L. G., Ross, J. A. (1992), Recombination at clean and decorated misfit dislocations. *Applied Physics Letters*, **60**, 1096–1098.

Raviswaran, A., Liu, C.-P., Kim, J., Cahill, D. G., Gibson, M. J. (2001), Evolution of coherent islands during strained-layer Volmer-Weber growth of Si on Ge(111). *Physical Review B*, **63**, 125314+.

Schaffler, F., Tobben, D., Herzog, Abstreiter, G., Holländer, B. (1992), High-electron-mobility Si/SiGe heterostructures: influence of the relaxed SiGe buffer layer. *Semiconductor Science and Technology*, **7**, 260–266.

Shimizu, T., Zhang, Z., Shingubara, S., Senz, S., Gösele, U. (2009), Vertical epitaxial wire-on-wire growth of Ge/Si on Si(100) substrate. *Nano Letters*, **9**, 1523–1526.

Shiryaev, S.Y., Jensen, F., Hansen, J. L., Petersen, J. W., Larsen, A. N. (1997), Nanoscale structuring by misfit dislocations in $Si_{1-x}Ge_x$/Si epitaxial systems. *Physical Review Letters*, **78**, 503–506.

Sze, S. M. (1981), *Physics of Semiconductor Devices*, New York, John Wiley & Sons.

Takagi, S. (2003), Re-examination of subband structure engineering in ultra-short channel MOSFETs under ballistic carrier transport. *Symposium on VLSI Technology Digest of Technical Papers,* 115–116.

Tersoff, J., LeGoues, F. K. (1994), Competing relaxation mechanisms in strained layers. *Physical Review Letters*, **72**, 3570–3573.

Thompson, N. (1953), Dislocation nodes in face-centred cubic lattices. *Proceedings of the Physical Society, Section B*, **66**, 481–492.

Tuppen, C. G., Gibbings, C. J., Hockly, M. (1991), Low threading dislocation densities in thick, relaxed $Si_{1-x}Ge_x$ buffer layers. In Bean, J. C., Iyer, S. S. and Wang, K. L. (eds), *Mater. Res. Soc. Symp. Proc.*, **220**, 187–192.

Tyagi, M. S., Van Overstraeten, R. (1983), Minority carrier recombination in heavily-doped silicon. *Solid-State Electronics*, **26**, 577–597.

Verlinden, P., Evrard, O., Mazy, E., Crahay, A. (1992), The surface texturization of solar cells: A new method using V-grooves with controllable sidewall angles. *Solar Energy Materials and Solar Cells*, **26**, 71–78.

Vernon, S. M., Tobin, S. P., Al-Jassim, M. M., Ahrenkiel, R. K., Jones, K. M., Keyes, B. M. (1990), Experimental study of solar cell performance versus dislocation density. *Photovoltaic Specialists Conference, Conference Record of the Twenty First IEEE*, 211–216.

Wagner, G. (1998), Defect structure of strained heteroepitaxial $In_{1-x}Al_xP$ layers deposited by MOVPE on (001) GaAs substrates. *Crystal Research and Technology*, **3**, 383–400.

Wegscheider, W., Eberl, K., Abstreiter, G., Cerva, H., Oppolzer, H. (1990), Novel relaxation process in strained Si/Ge superlattices grown on Ge (001). *Applied Physics Letters*, **57**, 1496–1498.

Wu, Y., Fan, R., Yang, P. (2002), Block-by-block growth of single-crystalline Si/SiGe superlattice nanowires. *Nano Letters*, **2**, 83–86.

Xiang, J., Lu, W., Hu, Y., Wu, Y., Yan, H., Lieber, C. M. (2006), Ge/Si nanowire heterostructures as high-performance field-effect transistors. *Nature*, **441**(7092), 489–493.

Xie, Y. H., Gilmer, G. H., Roland, C., Silverman, P. J., Buratto, S. K., Cheng, J. Y., Fitzgerald, E. A., Kortan, A. R., Schuppler, S., Marcus, M. A., Citrin, P. H. (1994), Semiconductor surface roughness: Dependence on sign and magnitude of bulk strain. *Physical Review Letters*, **73**, 3006–3009.

Xie, Y. H., Samavedam, S. B., Bulsara, M., Langdo, T. A., Fitzgerald, E. A. (1997), Relaxed template for fabricating regularly distributed quantum dot arrays. *Applied Physics Letters*, **71**, 3567–3568.

Yang, M., Chan, V. W. C., Chan, K. K., Shi, L., Fried, D. M., Stathis, J. H., Chou, A. I., Gusev, E., Ott, J. A., Burns, L. E., Fischetti, M. V., Ieong, M. (2006), Hybrid-orientation technology (HOT): opportunities and challenges. *IEEE Transactions on Electron Devices*, **53**, 965–978.

Zeng, L., Yi, Y., Hong, C., Liu, J., Feng, N., Duan, X., Kimerling, L. C., Alamariu, B. A. (2006), Efficiency enhancement in Si solar cells by textured photonic crystal back reflector. *Applied Physics Letters*, **89**, 111111.

14

Transport properties of silicon/silicon–germanium (Si/SiGe) nanostructures at low temperatures

A. GOLD, Université Paul Sabatier, France

Abstract: We present theoretical results on the transport properties of SiGe/Si/SiGe quantum wells and Si/SiGe heterostructures at low temperatures. Charged-impurity scattering and interface-roughness scattering are considered. Mobility limits are discussed. For different scattering mechanisms the transport scattering time and the single-particle scattering time are calculated and compared. We also study the effects of an in-plane magnetic field applied to the two-dimensional electron gas. The transport properties in such spin-polarized electron gases can give information about the disorder. We describe different kind of measurements in order to get information about microscopic parameters of disorder. We present a discussion of the disorder-induced metal–insulator transition. Some predictions concerning the disorder-induced metal–insulator transition in Si/SiGe nanostructures are made.

Key words: transport scattering time, metal–insulator transition, single-particle relaxation time, Si/SiGe quantum well, Si/SiGe heterostructures.

14.1 Introduction

Transport properties in silicon MOSFET structures[1] and GaAs/AlGaAs heterostructures and quantum wells[1,2] have been studied for many years and are still under research. The electron density in MOSFET structures can be modified with a gate. The mobility in such structures is limited by impurities at the Si/SiO$_2$ interface and by interface-roughness scattering.[1] The peak mobility[3] at low temperature is around $\mu \approx 3 \times 10^4$ cm^2/Vs in high-mobility samples, which is a low mobility value compared to that of GaAs/AlGaAs heterostructures. Due to remote doping, the mobility in GaAs/AlGaAs structures is higher and $\mu \approx 8 \times 10^6$ cm^2/Vs has been reached.[2] For a long time the variation of the electron density in GaAs/AlGaAs structures with a gate was difficult to obtain. Therefore, in publications one could not present results as a function of density for one sample with a given amount of disorder. Often density was changed by light.

In Si$_{1-x}$Ge$_x$/Si/Si$_{1-x}$Ge$_x$ quantum wells (QWs) and Si/Si$_{1-x}$Ge$_x$ heterostructures (HSs), where the two-dimensional electron gas is located in the silicon, the situation is similar to the case of GaAs/Al$_x$Ga$_{1-x}$As QWs and HSs. The first samples for a two-dimensional electron gas showed a low mobility of $\mu < 2.4 \times 10^3$ cm^2/Vs.[4] Transport properties in QW structures have been studied theoretically for low temperatures, and mobility limits have been discussed from a theoretical point of view.[5] Heterostructures are very similar to MOSFET structures.[3] During the last 25 years the low-temperature mobility in such Si/Si$_{1-x}$Ge$_x$ samples has increased considerably.[6,7] A recent high-mobility sample in a HS showed a peak mobility of about $\mu \approx 1.6 \times 10^6$ cm^2/Vs.[8] Only in recent years has it become possible to modify the electron density with a gate. However, the density variation is still relatively small, around a factor of 2–3, and only very recently has one been able to obtain a value within one order of magnitude: 2.5×10^{10} cm^{-2} $< N < 3 \times 10^{11}$ cm^{-2}.[8]

Due to the absence of a gate in early experimental work and the limiting density range using a gate in recent work, no systematic measurements have been performed that allow one to understand in detail the parameters of disorder present in such Si/Si$_{1-x}$Ge$_x$ structures. This is particularly true for the parameters of interface-roughness scattering (IRS). It is generally believed that impurity scattering (IS) is dominant at lower electron density and IRS is important at higher electron density if the QW width is thin enough.[5] In heterostructures it is clear that IRS will be important at high density. This is similar to the situation in silicon MOSFET structures.[1] However, parameters for IRS have not been convincingly determined. For IS the doping profile really present in a sample has been determined by some measurements. We believe that the reason for the lack of systematic work is the limiting range of electron density available in experiment. To compare experiment with theory a larger density range must be available.

In general one can say that there are few systematic studies in Si/SiGe. Experimenters are mainly interested in either (i) high-mobility samples for applications or (ii) samples with low carrier density for research. For low carrier density, many-body effects (exchange and correlation) are important. However, at low carrier density it is difficult to compare experimental results with theory because either (i) a metal–insulator transition (MIT) occurs, or (ii) a new ground state is observed, for instance a Wigner crystal. For new ground states, besides the Fermi-liquid ground state (see Section 14.2), trustable and confirmed theories concerning the transport properties are not available from the literature. Consequently, one can find many unconfirmed predictions and speculations in the literature.

In order to reduce disorder in future samples and to get higher mobility, one should try to understand the disorder present in existing samples and to determine the parameters that characterize this disorder in connection with how the sample was produced. More importantly, in future nanostructures, which will become smaller and smaller, one should know if the doping profile, intended in the production process, corresponds to the real doping profile present in the real structure. Mobility measurements help to get microscopic insight.

In the present chapter we shall describe transport measurements that help to get microscopic insight into the source of disorder (IS and IRS). We present calculations of transport properties of the two-dimensional electron gas in QWs and HSs made at low temperatures and we point out the characteristic behaviour for IS and IRS. We believe that new ground states depend on the kind of disorder present in the samples, and that the knowledge of this disorder is an important ingredient for developing a theory of these new ground states. Also for technological issues the detailed knowledge of disorder is extremely important, especially when systems become smaller.

The chapter is organized as follows. In Section 14.2 we describe the model and the theory. The different scattering times are discussed and the behaviour in a parallel magnetic field and at low temperatures is described. The disorder-induced MIT is described. Theoretical results for transport properties of quantum wells are discussed in Section 14.3 and heterostructures are considered in Section 14.4. Our discussion of the theoretical results in comparison with transport experiments using Si/SiGe nanostructures is in Section 14.5. A general discussion with future trends is given in Section 14.6. We conclude in Section 14.7.

14.2 Model, disorder and transport theory

In this chapter we describe the model of the electron gas, used for QWs and HSs. We specify the disorder and we give some account of the transport theory used in the calculations. For more detailed information we advise the reader to consult the original literature.

14.2.1 Model

We consider an interacting two-dimensional electron gas in the *xy*-plane with an in-plane effective mass $m^* = 0.19\, m_e$ and confinement in the *z*-direction. m_e represents the free electron mass. The effective mass perpendicular to the xy-plane is $m_z = 0.916\, m_e$, m_z being important for IRS in the case of QWs and for confinement effects in the case of HSs. We consider QWs with infinite barriers at $z = 0$ and $z = L$ and the electron gas is located at $0 \leq z \leq L$ with a wave function in the *z*-direction given by $\xi_0(z) \propto \sin(\pi z/L)$.[5] The parameter in such a system is the QW width L and results for the mobility will depend on L. For Si/SiGe heterostructures we use the triangular confinement potential where the effective penetration z_0 of the Fang–Howard wave function $\xi_0(z < 0) \propto z \exp(bz/2)$ into the semiconductor ($z < 0$) is given by $z_0 = -3/b$ and b is the Stern and Howard confinement parameter.[1] $b > 0$ is determined by $b = (48\pi m_z e^2 N^*/\varepsilon_{sc}\hbar^2)^{1/3}$ with $N^* = N_{\mathrm{Depl}} + 11N/32$, where N is the electron density. For the depletion density N_{Depl} we use $N_{\mathrm{Depl}} = 0$ in the following. $\varepsilon_{\mathrm{sc}}$ represents the dielectric constant of the semiconductor, in our case of Si, and we use $\varepsilon_{\mathrm{sc}} = 12.5$. For the insulator SiGe we use $\varepsilon_{\mathrm{is}} = 12.5$. In the case of Si/SiGe heterostructures the width of the electron gas is determined by b and is not an independent parameter. In QWs the width L is an independent parameter.

We take into account Coulomb interaction effects between the electrons within the random-phase approximation (RPA),[9] modified by a local-field correction (LFC) $G(q)$. This leads to screening of the random potential by a q-dependent dielectric function[1,10]

$$\varepsilon(q) = 1 + q_{\mathrm{S}}[1 - G(q)]F_{\mathrm{c}}(q)/q \tag{14.1}$$

where $q_{\mathrm{S}} = g_{\mathrm{s}}g_{\mathrm{v}}/a_{\mathrm{B}}^*$ is the screening wave number given by the valley degeneracy $g_{\mathrm{v}} = 2$, the spin degeneracy $g_{\mathrm{s}} = 2$, and the effective Bohr radius $a_{\mathrm{B}}^* = a_0\varepsilon_{\mathrm{L}}m_{\mathrm{e}}/m^*$, defined by the effective mass m^* and the Bohr radius of the hydrogen atom $a_0 = 0.529$ Å. $\varepsilon_{\mathrm{L}} = (\varepsilon_{\mathrm{sc}} + \varepsilon_{\mathrm{is}})/2$ represents the dielectric constant of the background material. For Si/SiGe we use $\varepsilon_{\mathrm{L}} = 12.5$ and $a_{\mathrm{B}}^* = 34.8$ Å. $F_{\mathrm{c}}(q)$ represents the form factor for the Fourier-transformed Coulomb interaction potential $V(q) = 2\pi e^2F_{\mathrm{c}}(q)/\varepsilon_{\mathrm{L}}q$ of the two-dimensional electron gas. Expressions for $F_{\mathrm{c}}(q)$ for HSs can be found in Ref. 1 and for QWs in Ref. 5. For QWs $F_{\mathrm{c}}(q)$ depends on the width L and for MOSFET structures it depends on the width b. In the long-wave length limit one finds that $F_{\mathrm{c}}(q \to 0) = 1$.

The LFC takes into account many-body effects (exchange and correlation) that are neglected in the RPA.[9,10] A finite LFC leads to reduced screening and therefore also to reduced scattering times (reduced mobility), if compared to the RPA. These many-body effects can be neglected for high electron densities, where the RPA is exact. However, at lower electron density the LFC plays an important role. Analytical expressions obtained from a fit of the LFC as calculated[11,12] for a two-dimensional electron gas are used in the present chapter. We stress that the valley and spin degeneracy is taken into account in these calculations of the LFC. This is a very important point.

Many silicon nanostructures are made from SiGe/Si/SiGe QWs. Then, from a principal point of view all structures should have a QW confinement. If doping is symmetric with doping left and right of the QW, this model of confinement is completely justified. In the case of non-symmetric doping and/or the application of an electric field using a gate, it might be better to describe the confinement by a HS as for Si/SiGe HSs. In this case many of the results obtained for MOSFET structures are applicable, at least qualitatively, because only the background dielectric constant is changed, from 7.7 in silicon MOSFET to 12.5 in Si/SiGe. This change of the dielectric constant improves the mobility in Si/SiGe due to better screening and a larger effective Bohr radius compared to silicon MOSFET structures. Because of this the Wigner–Seitz parameter r_s becomes, for given density, smaller in SiGe compared to MOSFET.

14.2.2 Disorder

As the relevant source of disorder we consider IS and IRS. IS represents disorder due to randomly distributed charged impurities of density N_i in the *xy*-plane at a distance z_i from the interface. For QWs $z_i = L/2$ corresponds to impurities located at the centre of the QW, and for $z_i = 0$ (or $z_i = L$) the impurities are located at the edge of the QW. For $z_i < 0$ the QW is remote doped. In HSs $z_i = 0$ corresponds to impurities located at the interface of the semiconductor (Si) and the insulator (SiGe). For $z_i < 0$ the impurities are located in the region of the electron gas and $z_i > 0$ represents remote impurities. The Fourier transform of the unscreened random potential for IS is written as[1]

$$\langle |U(q)|^2 \rangle = N_i [2\pi e^2 F_i(q, z_i)/\varepsilon_L q]^2 \qquad 14.2$$

The factor $F_i(q, z_i)$ represents the impurity form factor, which also depends on L if QWs are considered. In the case of heterostructures $F_i(q, z_i)$ depends on b. Explicit expressions can be found in Ref. 1 for HS and in Ref. 5 for QWs. We recall that in the long wavelength limit one gets $F_i(q = 0, z_i) = 1$.

In Ref. 5 we have discussed homogeneous background doping. In adding doping layers of different z_i one can create doping of the volume. In this chapter we do not discuss this scattering mechanism because it is not possible to keep the two-dimensional density of the doping charge finite and the global neutrality condition cannot be satisfied.

We also consider IRS, described by Δ, corresponding to the average height of the roughness in the *z*-direction, and Λ, corresponding to the correlation length parameter of the roughness in the *xy*-plane. One expects $\Lambda >> \Delta$ and Δ should be a multiple of the distance between neighbouring planes in Si. We use Λ as a fit parameter with $\Delta \approx 3$Å. The correlation in the *xy*-plane is assumed to be Gaussian. For QWs the random potential is given by[13]

$$\langle |U(q)|^2 \rangle = 2\pi^5 \Delta^2 \Lambda^2 \exp(-q^2 \Lambda^2/4)/(L^6 m_z^2) \qquad 14.3$$

The factor 2 represents the two interfaces of the QW. It was *predicted* that in thin QWs IRS is dominant and the mobility μ increases strongly as $\mu \propto L^6$ with increasing width.[13] This prediction of IRS in thin quantum wells was *verified* in experiments by mobility measurements in GaAs QWs[14,15] and In As QWs.[16] Note that the large value of m_z decreases IRS in silicon QWs.[5] For IRS[17] in HS we use the expression[1,18]

$$\langle |U(q)|^2 \rangle = 4\pi^3 e^4 \Delta^2 \Lambda^2 (N + 2N_{\mathrm{Depl}})^2 \exp(-q^2 \Lambda^2/4) \varepsilon_L^2 \qquad 14.4$$

Because $\langle |U(\vec{q})|^2 \rangle \propto N^2$ IRS becomes dominant at high electron density where the electron gas is pushed to the interface and the disorder due to IRS increases strongly, as in the case of MOSFET structures.[1]

14.2.3 Scattering times

In the following we present calculations of the mobility μ_0 in the lowest order $\langle |U(\vec{q})|^2 \rangle > 0$ of the random potential for the disorder. The 0 indicates that the mobility is calculated in the lowest order of the random potential. We assume that $\langle U(\vec{q}) \rangle \equiv 0$. The mobility is given in terms of the transport scattering time $\tau_t^{(0)}$ via $\mu_0 = e\tau_t^{(0)}/m^*$. The inverse transport scattering time is expressed as[19]

$$\frac{\hbar}{\tau_t^{(0)}} = \frac{1}{2\pi\varepsilon_F} \int_0^{2k_F} \mathrm{d}q \, \frac{q^2}{\sqrt{4k_F^2 - q^2}} \frac{\langle |U(q)|^2 \rangle}{\varepsilon(q)^2} \tag{14.5}$$

In this formula the Fermi energy ε_F and the Fermi wave number k_F enter. Note that for IS $\mu_0 \propto 1/N_i$ and the numerical results shown in the figures can be used to obtain real numbers for the impurity concentration in real samples. However, the position of impurities must be known. For IRS with $\mu_0 \propto f(\Lambda)/\Lambda^2\Delta^2$ the dependence on Λ is more involved, as will be discussed later.

The single-particle relaxation time $\tau_s^{(0)}$ is accessible by measuring Shubnikov–de Haas oscillations in a small magnetic field oriented perpendicular to the two-dimensional electron gas.[1] From a theoretical point of view this time is related to the density of states correction due to disorder. From an experimental point of view it is related to the Dingle temperature[1] $T_D = \hbar/(2\pi\tau_s k_B)$ and expressed as[19]

$$\frac{\hbar}{\tau_s^{(0)}} = \frac{1}{\pi\varepsilon_F} \int_0^{2k_F} \mathrm{d}q \, \frac{k_F^2}{\sqrt{4k_F^2 - q^2}} \frac{\langle |U(q)|^2 \rangle}{\varepsilon(q)^2} \tag{14.6}$$

It is clear that $\tau_s^{(0)} \propto \tau_t^{(0)}$. However, there exists a characteristic difference. For the transport scattering time, back-scattering with $\Delta q \equiv q_f - q_i \approx 2k_F$ is most important, while for the single-particle relaxation time all scattering events are important. This leads, for an ideally two-dimensional electron gas with no extension in the z-direction, to the following relation in the case of IS[19]

$$1 \leq \tau_t^{(0)}/\tau_s^{(0)} \leq (2k_F|z_i|)^2 \tag{14.7}$$

For remote doping one finds that for large electron density $2k_F$–scattering (large angle scattering), which contributes strongly to the resistance and therefore to the inverse transport time, is strongly reduced.[20] Therefore, the scattering time is strongly enhanced and $\tau_t^{(0)}/\tau_s^{(0)} >> 1$. For IRS one finds[19]

$$2/3 \leq \tau_t^{(0)}/\tau_s^{(0)} \leq (k_F\Lambda)^2/3 \qquad 14.8$$

The lower limit in Eqs 14.7 and 14.8 is for low electron density and the limit $\tau_t^{(0)}/\tau_s^{(0)} \propto k_F^2 \propto N$ is for high electron density. Measuring both scattering times at high electron density gives information about z_i in the case of IS and about Λ in the case of IRS. At low density a value of $\tau_t^{(0)}/\tau_s^{(0)} \approx \frac{2}{3}$ indicates the importance of IRS. It should be noted that this only holds in the lowest order. For more results near a MIT see Ref. 19, where it was predicted that it might be possible to observe Shubnikov–de Haas oscillations with a finite $\tau_s^{(0)} \propto 1/N_i$ near and in the insulator phase.

14.2.4 Spin-polarized electron gas

By applying a parallel magnetic field to a two-dimensional electron gas one can spin-polarize the carriers.[21] The magnetic field B for complete spin polarization is given by $B > B_c = 2\varepsilon_F/g^*\mu_B = 2hN/(eg_v g^* m^*/m_e)$ where g^* is the effective Landé g-factor and μ_B is Bohr's magneton. In thin systems such as silicon MOSFET structures and SiGe HSs orbital effects[22] can be neglected. It was shown[21] that in this case the magnetic field acts on the spin degeneracy. A transport theory was developed for a partially[23,24] and fully polarized[24,25] two-dimensional electron gas in the presence of different scattering mechanisms. The transport theory takes into account the modifications of the density of states by a parallel magnetic field, the modification of the screening wave number, and the change of the Fermi wave number in the interacting electron gas. Note that screening is weaker, $\varepsilon(q) \propto q_s/q \propto g_v g_s/q$, in a spin-polarized system and consequently the screened random potential $\langle|U(\vec{q})|^2\rangle/\varepsilon(q)^2$ is larger than in a non-polarized electron gas. Of course, the Fermi energy is larger by a factor of 2 in the spin-polarized system. As far as the spin is concerned, this means that for zero magnetic field one has a non-polarized electron gas with $g_s = 2$ characterized by a resistance $\rho_0(B = 0) \equiv 1/(eN\mu_0)$; for $0 < B < B_c$ the resistance increases and it becomes constant, $\rho_0(B > B_c) = \rho_0(B_c)$, for $B > B_c$.[23] For $B > B_c$ the electron gas is completely spin-polarized with $g_s = 1$. The screening properties of a spin-polarized electron gas are strongly reduced compared to the non-polarized electron gas.

Neglecting the LFC and finite width effects of the electron gas and for impurities in the plane of the electron gas, one gets[23]

$$1/2 < \rho_0(B_c)/\rho_0(B = 0) < 4 \qquad 14.9$$

For impurities separated by a spacer the resistance ratio at intermediate electron densities is reduced,[25] $\rho_0(B_c)/\rho_0(B = 0) \approx 1.5$, compared to the case where impurities are located in the plane of the electron gas. For a short-range random potential, as for IRS, neglecting the exponential cut-off of the random potential at large Fermi wave numbers, one finds[24]

$$1 < \rho_0(B_c)/\rho_0(B = 0) < 8 \qquad 14.10$$

The strong increase of the resistance ratio in Eqs 14.9 and 14.10 occurs at low electron density. We suggest that one can use transport measurements to get information on the scattering parameters, especially in doing experiments at low carrier density, where the critical field for complete spin-polarization is small: $B_c \propto N$.

14.2.5 Temperature dependence of the mobility

Our theoretical results described before are obtained for zero temperature. For low temperature an approximately linear temperature dependence of the mobility for IS was found numerically.[26] This was not expected in a conventional Sommerfeld expansion, where a T^2 appears. A linear temperature dependence of the mobility was proven analytically[27] and was traced back to the anomalous wave-number dependence of the screening function (or the compressibility). The origin of this effect is the behaviour of the compressibility of the free electron gas in two dimensions as a function of the wave number q, also responsible for Friedel oscillations, which is of course well known.

For low temperatures the analytical result for the mobility was given as[27]

$$\mu_0(T) = \mu_0\left(1 - C(\alpha, N)\,\frac{k_B T}{\varepsilon_F} + O(T^{3/2})\right) \qquad 14.11$$

The coefficient $C(\alpha, N) > 0$ depends on the kind of disorder ($\alpha = 0$ for IRS and $\alpha = -1$ for IS), the carrier density and the LFC. Analytical and numerical results for $C(\alpha, N)$ have been given and discussed in Ref. 28 by using the Hubbard approximation for the LFC. Later it was shown that for remote doping with large spacer the coefficient $C(-1, N)$ becomes exponentially small.[29] Spin-polarized electron gases also have been discussed including the LFC for $g_s = 1$.[30] We stress that in the numerical result[26] and in the analytical result[27] a factor of 2 is missing. This factor of 2 was corrected in Ref. 28. For IS the factor of 2 is clearly seen in Table 1 of Ref. 28. We mention that at low density the LFC becomes very important[28] and it is the validity range of the LFC that determines the accuracy of Eq. 14.11. The form factor $C(-1, N)$ for Si/SiGe QWs has already been presented and discussed in Ref. 5. We also note that it is conceivable that $C(\alpha, N) < 0$.[30] This could happen in a strongly correlated electron gas when the LFC fulfils $G(q \approx 2k_F) > 1$ and $F(2k_F)[G(q \approx 2k_F) - 1] < 2k_F/q_s$. We predict that this can happen more easily in GaAs HSs with $g_v = 1$ than in Si/SiGe systems with $g_v = 2$.

The linear temperature dependence of the mobility was confirmed within a diagrammatic approach[31] and expressed in terms of an unknown Fermi-liquid parameter. Therefore we argued before[30] that the diagrammatic approach has no predictive power. The diagrammatic approach uses a short-range random potential as disorder. Therefore, not even the conductivity at zero temperature, calculated within this approach, can be compared to experiments. We stress that for mobility calculations of QWs and HSs, form factors for extension effects and the LFC enter the calculation. These form factors lead, even for IRS, to a non-short-ranged random potential. For IS the random potential is already not short-ranged from the beginning because of the q-dependence. Moreover, the diagrammatic approach is a perturbation approach and the temperature-dependent corrections must be small compared to the result at zero temperature. The T-dependent screening approach of Refs 26–30 is not a perturbation approach and the temperature-dependent part of the mobility could, in principle, be much larger than the mobility at zero temperature. However, we admit that for obtaining Eq. 14.11, an approximation was used: it was assumed that the corrections due to a finite temperature are small.[28] We add that the RPA+LFC approach is not limited to small r_s values: the use of the LFC allows one also to consider strongly correlated systems. In conclusion, we would like to ask the question: why should one use (and trust) a diagrammatic approach parameterized by an unknown Fermi-liquid parameter when this approach cannot describe the conductivity at zero temperature? We insist that the calculation in Ref. 27 was the first calculation providing an anomalous (linear) T-dependence of physical properties in two dimensions. Much later, many people used this idea.

The diagrammatic approach was widely used by experimenters in order to determine the Fermi-liquid parameter. Most experiments have been made with samples where IS was important, for which, however, the diagrammatic approach cannot be used, because the random potential is not short-ranged. We find it very surprising that it is so exciting to some experimenters to compare their work to a theory having unknown parameters,[31] instead of comparing it with a theory having known parameters[28] and, therefore, predictive power.

14.2.6 Multiple-scattering effects

Suppose that one increases, for given electron density, the disorder. Then multiple-scattering effects become important and the lowest-order theory in disorder, used until now, can no longer be valid. Experimentally it was shown that a metal-insulator transition (MIT) occurs in silicon MOSFET structures when the electron density is reduced – see the experiments discussed in Refs 1–3. The MIT occurs at a critical electron density (the mobility edge), with a metallic phase at high electron density and an insulating phase at low density. Unfortunately, these old experiments, showing a disorder-induced MIT, are not discussed in reviews on the MIT.[32] Multiple-scattering effects lead to a disorder-driven MIT. Within a self-consistent mode-coupling approach it was shown that for a non-interacting electron gas[33] and for an interacting electron gas in two dimensions, a MIT occurs.[34] In the mode-coupling approach weak localization effects[35] are not taken into account, but are neglected. This is justified: it was shown in an experiment using silicon MOSFET that weak localization effects do not lead to a MIT.[36]

In the following we apply for the mobility an expression obtained for a non-interacting electron gas,[37] generalized to an interacting electron gas.[38] With multiple-scattering effects taken into account, the mobility μ is given by

$$\mu = \mu_0 \,(1 - A) \qquad 14.12$$

where it is assumed that $0 < A \leq 1$, which corresponds to the metallic phase. μ_0 represents the lowest-order result discussed before and for $A << 1$ multiple-scattering effects can be neglected, $\mu \approx \mu_0$. For $A > 1$ the mobility at zero temperature vanishes: $\mu = 0$, which represents the insulating phase. Near the critical density the mobility will vanish as $\mu \propto (N - N_c)$. But at this point we want to make clear that the mode-coupling theory is a mean-field type theory. Therefore it is not expected to get correct critical exponents. However, we claim that the mode-coupling theory can describe the mobility scale and the critical density to good account, compared to experiments. Using this approach one can interpret experimental results and one can make predictions.

The parameter A describes the importance of multiple-scattering effects and is related to the screened disorder by[33,34]

$$A = \frac{1}{4\pi N^2} \int_0^\infty q\mathrm{d}q \, \frac{\langle |U(q)|^2 \rangle}{\varepsilon(q)^2} \, g_0(q)^2 \qquad 14.13$$

where $g_0(q)$ represents the compressibility of the two-dimensional free electron gas.[1,33,34] The MIT occurs for $A = 1$, which defines the critical electron density $N_c = N_c(L, N_i, z_i, g_v, g_s)$ of the MIT. The metallic phase is characterized by $N > N_c$ and the insulating phase by $N < N_c$. Within this approach we can consider different kinds of disorder and can obtain explicit values for the phase diagram $N_c = N_c(N_i, z_i)$ in the case of IS or $N_c = N_c(L, \Delta, \Lambda)$ in the case of IRS. Of course, also two different kinds of disorder can be taken into account, $N_c = N_c(N_i, L)$. Even more important, also a spin-polarized electron gas can be considered and the effect of spin-polarization on the critical electron density was studied.[39]

We note here that in the mode-coupling theory one can describe weakly correlated systems, where the Wigner–Seitz parameter, defined in two dimensions as $r_s \equiv 1/\sqrt{\pi N a^{*2}}$, is small, $r_s < 1$. The range $r_s << 1$ represents the validity range of the RPA.[9] For Si/SiGe $r_s = 1$ corresponds to $N = 2.63 \times$

10^{12} cm^{-2}. Most experiments are made at small electron densities, where $3 < r_s < 10$. Strongly correlated electron systems, where exchange and correlation are strong and where $r_s > 1$, can also be described be the mode-coupling theory, but the results depend on the LFC used in the calculation. The use of a finite LFC becomes crucial. But this is well known, already for the clean system.[9]

14.2.7 Metal–insulator transition

In the past 25 years we have analysed many mobility measurements using the mode-coupling approach: silicon MOSFET structures,[3,19,33,34,39–42] Si/SiGe QWs,[5,25] GaAs HSs[38,43], and GaAs QWs[19] and InGaAs QWs.[19] The dynamic conductivity has been studied, too.[44,45] Good agreement between theory and experiment has been found. Below we give additional numerical results for the MIT in Si/SiGe QWs and HSs. The results for HSs are very similar to results for silicon MOSFET structures.[3,34] Presently studied Si/SiGe structures show a very high mobility[8] and the corresponding parameter range of such structures has not been studied before for MOSFET structures.

It is astonishing that our results concerning the MIT have not been considered in review papers on the MIT – see, for instance, Ref. 32 – especially since it is known that weak localization cannot lead to a MIT.[36] This might be related to the belief of the authors of Ref. 32 that the MIT is interaction induced and that this is the only existing MIT.[46] First, we don't believe that the MIT is interaction induced. We think that it is always disorder induced when there is disorder. We have already shown[41] that disorder is very important for the transport properties of the MIT occurring near a critical point where the effective mass is divergent. We stress that the mass divergence in silicon MOSFET structures[32,46] was discovered using the interpretation of the linear temperature dependence. The observation of the ferromagnetic instability[32] in silicon MOSFET structures, connected with the mass divergence, was measured by using the interpretation of the parallel magnetic field effect on the conductivity as described in Ref. 23. In the following we discuss the disorder-induced MIT in Si/SiGe nanostructures, which is defined by the condition $A = 1$.

We stress that the mode-coupling approach has a classic limit, discussed within the Lorentz model.[47] Therefore, percolation-like physics is included in the mode-coupling approach. However, when the transport scattering times depend on the Fermi energy, the Fermi wave number, the density of states or the spin degeneracy, then quantum effects are important. In this sense the mode-coupling approach is a quantum theory where weak localization effects are neglected. We stress that results concerning the MIT in systems with a very large spacer[38] do not depend on the Fermi wave number if $k_F z_i \to \infty$. Only for $k_F z_i \to \infty$ do the transport properties in such a system depend on the density and can be interpreted with the physics of percolation. For finite values of $k_F z_i$ the results depend weakly on k_F and quantum effects are important.[38]

We are interested in strong variations of the mobility with density, especially near the MIT,[38] and not in very weak temperature-dependent changes by logarithmic corrections due to weak localization effects in the metallic range.[35] We stress that the MIT occurring in the mode-coupling approach represents a quantum-phase transition induced by disorder. The MIT is obtained in a self-consistent calculation of the current–current correlation function in the spirit of the self-consistent Born approximation for the Green function.

In the mode-coupling approach, the effective mass is an input parameter and is assumed to be constant. There exist measurements of the effective mass in silicon MOSFET structures showing a diverging behaviour (at least a strong increase) near the critical density of the MIT.[32] For Si/SiGe structures no divergent effective mass has been reported in the literature. Temperature-dependent mobility data do not indicate such a divergent behaviour. Therefore we used a constant mass in our calculation. We think there are clear indications from experiments with Si/SiGe structures that a *disordered*-induced MIT exists. We shall describe this MIT in the following.

14.3 Transport in quantum wells

In our earlier paper[5] on transport properties of SiGe/Si/SiGe QWs published 23 years ago we used for the figures the relation $N_i = N$ in order to simulate changes of the electron density induced by light. In the present chapter we shall show all figures with fixed impurity density. We also use a finite LFC in our calculations. It will be shown that these effects are quite strong in the low-density range. Alloy-disorder scattering in QW was considered also in the literature but results in a very high mobility compared to experiment and is not yet important.[48] For transport of holes in QWs, see Ref. 49.

14.3.1 Lowest-order results for quantum wells

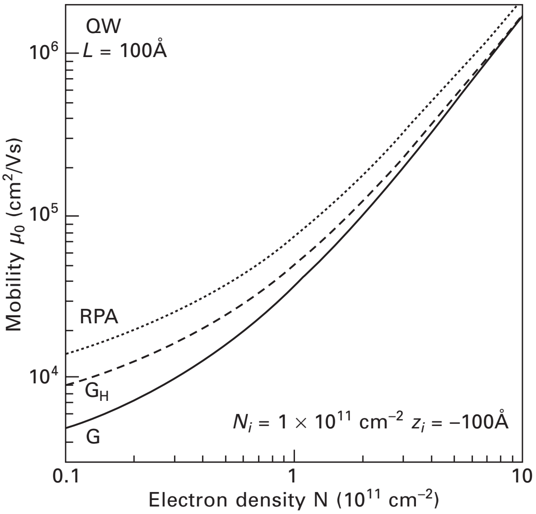
In Fig. 14.1 we show the mobility versus electron density for a QW of width $L = 100$ Å for IS ($N_i = 1 \times 10^{11}$ cm^{-2}, $z_i = -100$ Å) in different approximations for the LFC. A finite LFC reduce the mobility due to less screening. As can be seen, the differences become large at small densities, while at higher densities the differences can be neglected. At $N = 1 \times 10^{10}$ cm^{-2} the results of the full LFC and the RPA differ by a factor of nearly 5. This is not negligible! We stress that most theoretical calculations published in the literature neglect many-body effects and apply the RPA. In our work 23 years ago[5] we used the Hubbard approximation, where an analytical expression is available. In this approximation only exchange effects are taken into account and correlation effects are neglected. This approximation is also shown in Fig. 14.1.

The effect of remote doping is shown in Fig. 14.2 where the impurity distance z_i to the electron layer is increased from impurities in the centre of the QW ($z_i = L/2 = 50$ Å) to impurities at the edge of the quantum well ($z_i = 0$ Å) and then to impurities outside the QW ($z_i = -100$ Å and $z_i = -200$ Å). One can see a very large increase of the mobility due to remote doping, especially at high electron density where the Fermi wave number is large.

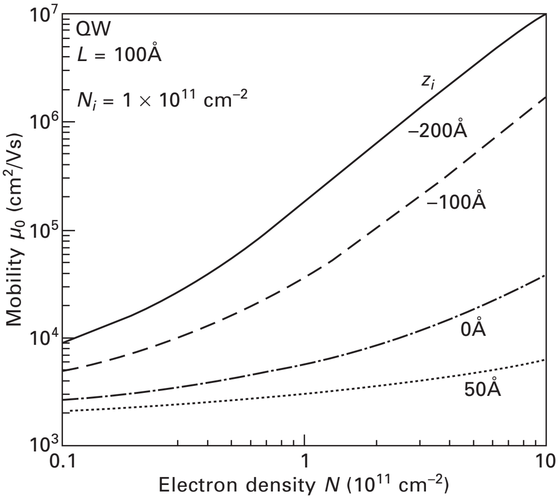
In Fig. 14.3 we show the mobility for IRS versus density for given IRS parameters ($\Delta = 3$Å, $\Lambda = 60$Å) for two well widths. Note that the two curves are related, to a very good approximation, by the law $\mu_0 \propto L^6$. This is because the form factor F_c(q) is only weakly L-dependent compared to the matrix element. At high density the mobility increases as $\mu_0 \propto \exp(k_F^2\Lambda^2)$ due to the exponential cut-off of the random potential for IRS.

The mobility versus density for different values of the IRS parameter Λ is shown in Fig. 14.4. With increasing Λ the minimum of the mobility at $N_{IRS}^* \approx 1/\pi\Lambda^2$ shifts to smaller density. Note that for N_{IRS}^* the mobility scales with $\mu_0(N_{IRS}^*) \propto 1/\Lambda$. It is very clear that for low density an IRS with a small value of Λ is favourable for getting high mobility.

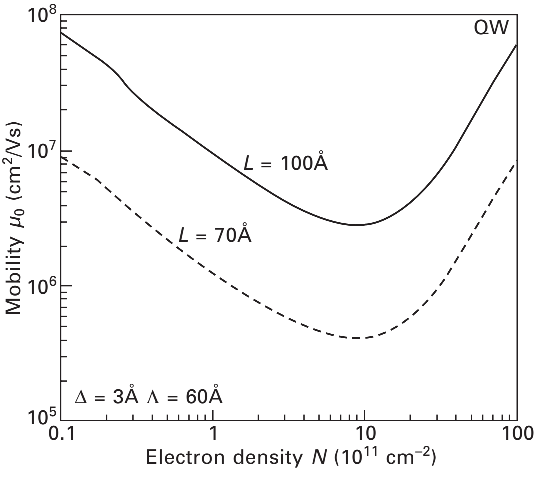
The single-particle relaxation time is evaluated in the following and compared to the transport scattering time. In Fig. 14.5 we show the ratio



14.1 Mobility μ_0 versus electron density N in a QW of width L = 100 Å with IS ($N_i = 1 \times 10^{11}$ cm^{-2}, $z_i = -100$ Å, which means impurities at 150 Å from the centre of the quantum well). Different expressions for the local-field correction are used in the calculations: the full form $G(q)$, the Hubbard approximation $G_H(q)$, and the RPA with $G(q) = 0$.

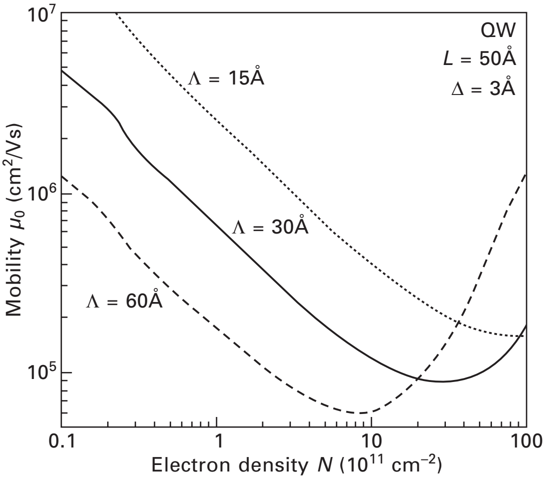


14.2 Mobility μ_0 versus electron density N for a QW of width L = 100 Å with IS ($N_i = 1 \times 10^{11}$ cm^{-2}) and different positions z_i of the impurity layer: z_i = –200 Å (solid line), z_i = –100 Å (dashed line), z_i = 0 Å (dashed-dotted line), and z_i = 50 Å (dotted line).

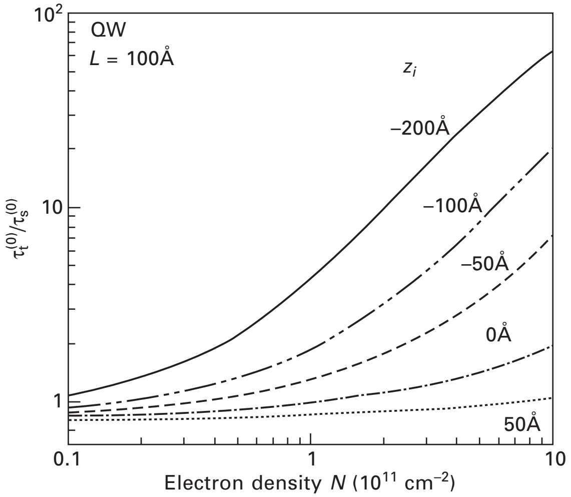


14.3 Mobility μ_0 versus electron density N for a QW with IRS of width L = 100 Å (solid line) and L = 70 Å (dashed line). The parameters for the IRS are Δ = 3 Å and Λ = 60 Å.

$\tau_t^{(0)}/\tau_s^{(0)}$ versus electron density for IS with different impurity positions. For remote doping the ratio increases strongly with density, while for impurities inside the QW a value near to $\tau_t/\tau_s \approx 1$ is found. For remote doping with $z_i < 0$ one expects $\tau_t^{(0)}/\tau_s^{(0)} \approx [2k_F(L/2 - z_i)]^2$. We therefore conclude that measurement of the ratio $\tau_t^{(0)}/\tau_s^{(0)}$ can give information about the impurity



14.4 Mobility μ_0 versus electron density N in a QW of width L = 50 Å with IRS (Δ = 3 Å). Different values for Λ are used: Λ = 60 Å (dashed line), Λ = 30 Å (solid line), and Λ = 15 Å (dotted line).



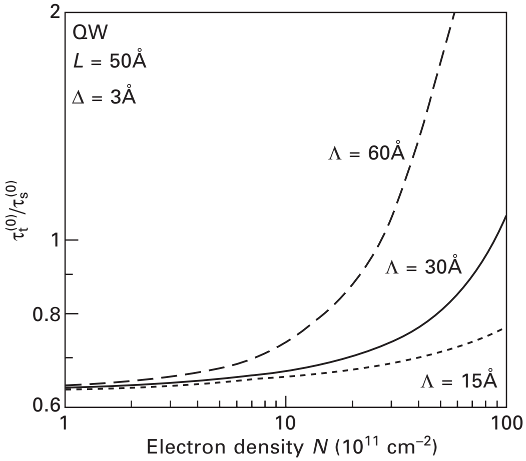
14.5 Scattering time ratio $\tau_t^{(0)}/\tau_s^{(0)}$ versus electron density N for a quantum well of width L = 100 Å with IS with different values of z_i: z_i = −200 Å (solid line), z_i = −100 Å (long-dashed line), z_i = −50 Å (short-dashed line), z_i = 0 Å (dashed-dotted line), and z_i = 50 Å (dotted line).

position. We note that this information is most easily obtained at a large electron density. For a low electron density one gets $\tau_t^{(0)}/\tau_s^{(0)} \approx 1$ within the RPA; however, a finite LFC leads to a value smaller than 1. This is visible in Fig. 14.5.

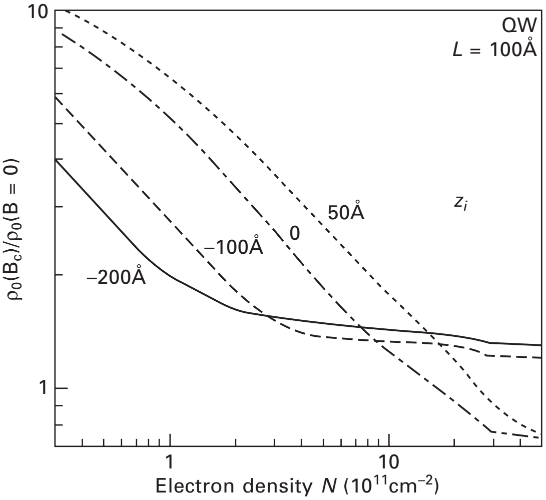
The ratio $\tau_t^{(0)}/\tau_s^{(0)}$ versus density is shown in Fig. 14.6 for IRS and different values of Λ. Note that for low density and within the RPA $\tau_t^{(0)}/\tau_s^{(0)} \approx 2/3$. A finite LFC leads to a small reduction. At higher density the ratio depends very much on the value of Λ. From Fig. 14.6 we can see that $\tau_t^{(0)}/\tau_s^{(0)} > 1$ for $N > 3N^*_{\mathrm{IRS}}$. Comparison of Figs 14.5 and 14.6 shows that $\tau_t^{(0)}/\tau_s^{(0)}$ can be used to get information about the scattering mechanism, especially at higher density, and to a lesser extent at low density.

With a parallel magnetic field the two-dimensional electron gas can be spin-polarized. The screening properties are reduced and one expects a larger resistance for the spin-polarized electron gas. We compare the resistance $\rho_0(B_c)$ of the spin-polarized electron gas with the resistance $\rho_0(B = 0)$ of the unpolarized electron gas. In Fig. 14.7 we show $\rho_0(B_c)/\rho_0(B = 0)$ versus electron density for a QW with IS for different impurity positions z_i. The ratio is strongly enhanced at low electron density, especially for impurities at the centre of the quantum well. For remote impurities a plateau region with $\rho_0(B_c)/\rho_0(B = 0) \approx 1.4$ can be seen at intermediate electron density. Within the RPA one finds in the low-density limit $\rho_0(B_c)/\rho_0(B = 0) = 4$.[23] As seen in Fig. 14.7, this limit behaviour is strongly increased by many-body effects via the LFC.[39]

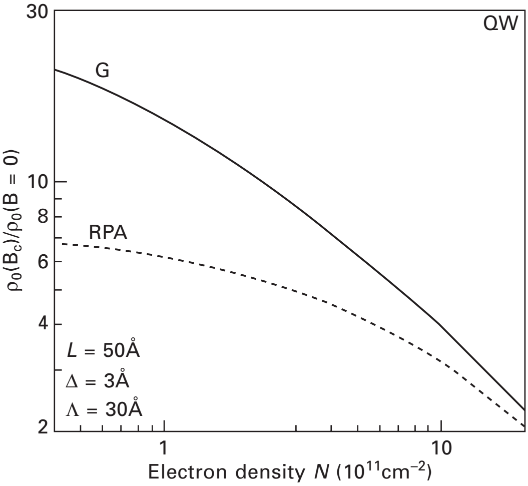
The effect of the LFC is demonstrated in Fig. 14.8 where we show $\rho_0(B_c)/\rho_0(B = 0)$ versus electron density for a QW with IRS. The width of Si/SiGe is so small that we expect that IRS is the dominant scattering mechanism. For IRS the limiting behaviour at low density within the RPA is $1 < \rho_0(B_c)/\rho_0(B$



14.6 Scattering time ratio $\tau_t^{(0)}/\tau_s^{(0)}$ versus electron density N for a quantum well of width L = 50 Å with IRS with Δ = 3 Å for different values of Λ: Λ = 15 Å (dotted line), Λ = 30 Å (solid line), and Λ = 60 Å (dashed line).



14.7 Resistance ratio $\rho_0(B = B_c)/\rho_0(B = 0)$ versus electron density N for a quantum well of width $L = 100$ Å with IS for different positions z_i of the impurity plane: $z_i = -200$ Å (solid line), $z_i = -100$ Å (dashed line), $z_i = 0$ Å (dashed-dotted line), and $z_i = 50$ Å (dotted line).



14.8 Resistance ratio $\rho_0(B = B_c)/\rho_0(B = 0)$ versus electron density N for a quantum well of width $L = 50$ Å for IRS ($\Delta = 3$ Å, $\Lambda = 30$ Å). The solid line was calculated with the LFC and the dotted line represents the RPA with $G(q) = 0$.

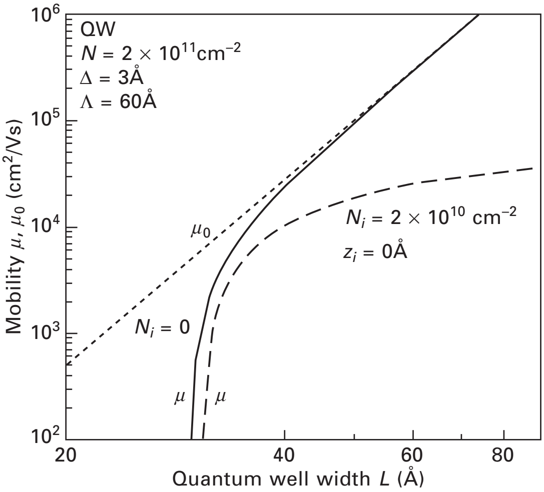
= 0) < 8, as visible by the dotted line. Many-body effects strongly increase this value. We believe that this effect could be used to get information about the LFC itself, especially at low electron density where the magnetic field for full spin polarization is small.

14.3.2 Metal–insulator transition in quantum wells

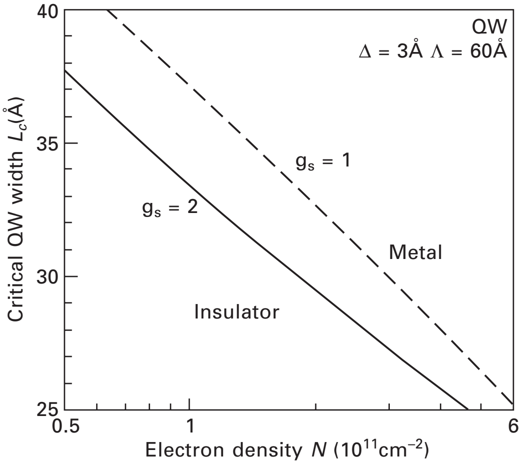
A disorder-induced MIT in QWs due to IRS was predicted in 1986.[13] Due to the strong increase of the disorder with decreasing well width, IRS becomes dominant in systems with small well width. Here we present some characteristic results obtained by taking into account multiple-scattering effects.

In Fig. 14.9 we show the mobility versus quantum well width for IRS. A MIT occurs at a critical width $L_c \equiv L \approx 30$ Å. Deviations of μ from the $\mu_0 \propto L^6$ law occur in a very restricted well width range of $L_c < L < 45$ Å $\approx 1.5L_c$. For the dashed line in Fig. 14.9 IRS and IS are taken into account. For large well width the mobility is determined by IS and becomes nearly constant as a function of the well width. Most important is the observation that $\mu_0 \approx L^6$ of the IRS cannot be seen due to the MIT occurring at small well width and the importance of IS at larger well width. Due to IS the critical well width L_c is slightly shifted to a larger value, see Fig. 14.9.

The critical QW width L_c versus electron density for IRS is shown in Fig. 14.10. L_c increases very slowly with decreasing electron density – note the logarithmic density range. The solid line is for a non-polarized electron gas with spin-degeneracy $g_s = 2$. The dashed line is for a fully spin-polarized electron gas with $g_s = 1$. For $L > L_c$ the system is metallic and for $L < L_c$ the system is insulating. With decreasing electron density, L_c increases: the insulating phase increases due to a factor $1/N^2$ in A. For a given density the



14.9 Mobility μ and μ_0 versus quantum well width L for electron density $N = 2 \times 10^{11}$ cm^{-2} for IRS (Δ = 3 Å, Λ = 60 Å). The dotted line represents the lowest-order result for IRS. For the solid and the dashed line multiple-scattering effects have been included. For the dashed line IRS and IS with z_i = 0 Å and $N_i = 2 \times 10^{10}$ cm^{-2} are taken into account (taken from Ref. 54).



14.10 Critical quantum well width L_c versus electron density for IRS ($\Delta = 3$ Å, $\Lambda = 60$ Å). The solid line is for the non-polarized electron gas and the dashed line is for the fully spin-polarized electron gas. $L > L_c$ corresponds to a metallic phase and $L < L_c$ to an insulating phase (taken from Ref. 54).

critical width is larger for the polarized system and the numerical factor is $L_c(g_s = 1)/L_c(g_s = 2) \approx 1.15$. For a fixed width the corresponding numerical factor is $N_c(g_s = 1)/N_c(g_s = 2) \approx 1.5$. This means that applying a parallel magnetic field to a non-polarized electron gas in the metallic phase can induce a MIT and the system is insulating when it is fully spin-polarized: this happens when the density and width of the system are between the solid and the dashed line shown in Fig. 14.10. The IRS induced MIT was predicted in Ref. 13. Results shown in Fig. 14.10 are similar to results obtained before for $g_s = 2$.[5] However, now also the spin-polarized case is discussed and the LFC is taken into account.

Finally we would like to note that $\tau_t/\tau_s^{(0)} < \tau_t^{(0)}/\tau_s^{(0)}$ and that at the MIT $\tau_t/\tau_s^{(0)} \to 0$. This is due to the fact that multiple-scattering effects only weakly influence τ_s but strongly reduce τ_t.[19] Measured ratios of $\tau_t/\tau_s < 0.5$ would indicate that one is near to a disorder-induced MIT. For transport measurements in a parallel magnetic field one can conclude that $\rho(B_c)/\rho(B = 0) > \rho_0(B_c)/\rho_0(B = 0)$ and that at the MIT $\rho(B_c)/\rho(B = 0) \to \infty$. This is due to the divergence of $\rho(B_c)$ at the MIT described by $N_c(g_s = 1) > N_c(g_s = 2)$ – see Fig. 14.10.[39]

14.4 Transport in heterostructures

In this section we report on the transport properties of Si/SiGe HSs with IS and IRS. Results are very similar to those for silicon MOSFET structures.[3,34]

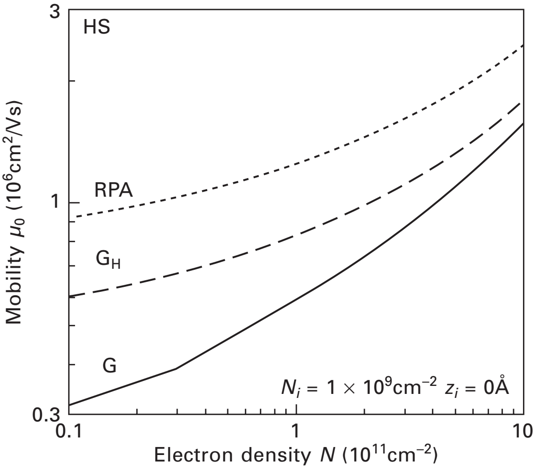
Theoretical results for Si/SiGe HSs and mobility limits, however, neglecting the LFC, can be found in the literature.[50,51] Transport properties of holes in heterostructures also have been calculated.[52]

For HSs we consider only impurities at the interface. The idea behind this is that the recent high-mobility sample[8] is undoped and the transport properties are possibly determined by *residual* impurities at the interface. The concentration of the charged impurities might be used to get agreement with experimental results. In most figures we show the interplay between IS and IRS. This helps to see clearly what kind of information one can get by studying transport properties in a larger density range.

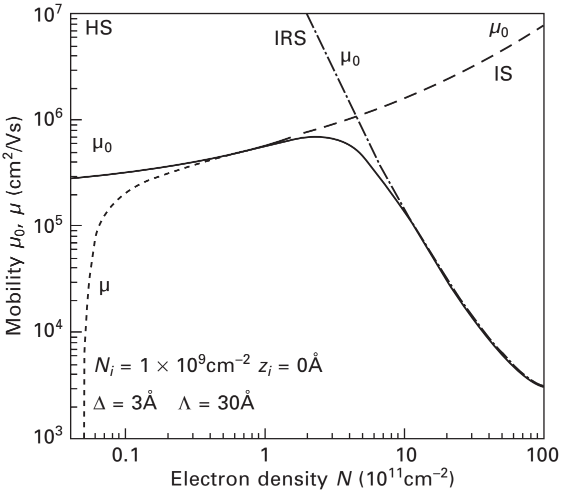
14.4.1 Lowest-order results for heterostructures

In Fig. 14.11 we show the influence of the LFC on the mobility for IS when shown versus electron density. Again, at low density the mobility, including the LFC, is much smaller compared to the RPA. Note that the increase of the mobility with density (the slope) is larger if the LFC is taken into account. Later, this will be important for the interpretation of an experiment. We stress that for impurities at the interface the mobility within RPA and within the Hubbard approximation becomes constant for vanishing electron density.

The mobility for IS and IRS is plotted versus density in Fig. 14.12. This shows the interplay between these two scattering mechanisms. The solid line is the calculation in lowest order of the disorder. A peak mobility $\mu_m \approx 7 \times 10^5 \mathrm{cm}^2\mathrm{Vs}$ is found at about $N_m \approx 3 \times 10^{11}\ \mathrm{cm}^{-2}$. Note that in order to



14.11 Mobility μ_0 versus electron density N in a heterostructure for IS with $N_i = 1 \times 10^9\ \mathrm{cm}^{-2}$ and $z_i = 0$ Å. Different forms of the local-field correction are used in the calculation: the full form $G(q)$, the Hubbard approximation $G_H(q)$, and the RPA with $G(q) = 0$.



14.12 Mobility μ_0 versus electron density N in a heterostructure with IS ($N_i = 1 \times 10^9$ cm^{-2}, $z_i = 0$ Å) and IRS ($\Delta = 3$ Å, $\Lambda = 30$ Å) as the solid line. Results for one scattering mechanism alone are shown as the dashed line (IS) and the dashed-dotted line (IRS). The dotted line represents μ where multiple-scattering effects are taken into account.

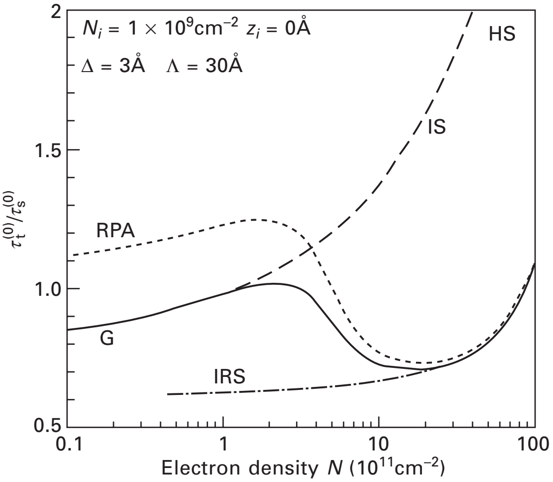
determine the parameters of the effects of IRS the electron density must be considerable larger than N_m.

For the same parameters as used in Fig. 14.12 we show in Fig. 14.13 the ratio $\tau_t^{(0)}/\tau_t^{(0)}$ versus density as the solid line. While at low density the ratio is determined by IS, at high electron density the ratio is determined by IRS. The transition regime is relatively sharp and helps to discriminate between IS and IRS. Figure 14.13 shows the importance of measurements of the transport scattering time and the single-particle scattering time. The dotted line was calculated with a vanishing LFC (RPA) and at low electron density one can see the effects of exchange-correlation by comparing with the solid line.

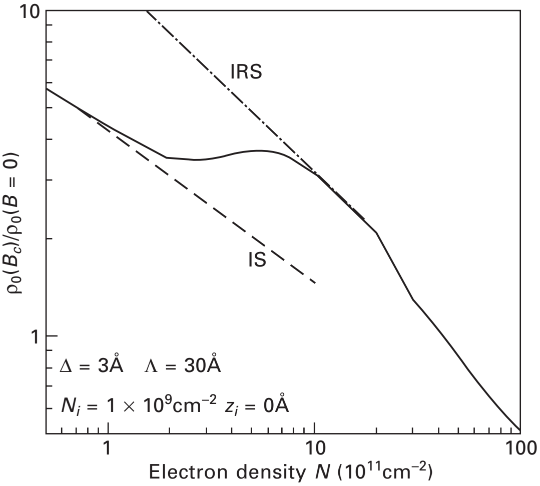
The magneto-resistance ratio $\rho_0(B_c)/\rho_0(B = 0)$ versus electron density for a HS with IS and IRS is shown in Fig. 14.14. Again we find a transition region between IS (dashed line) and IRS (dashed-dotted line) around the electron density of the peak-mobility. Note the strong enhancement $\rho_0(B_c)/\rho_0(B = 0) > 4$ at low density for IS due to the LFC. For IRS the ratio $\rho_0(B_c)/\rho_0(B = 0) > 8$ is also strongly enhanced at low density.

14.4.2 Metal–insulator transition in heterostructures

The dotted line in Fig. 14.12 represents the mobility due to IS and IRS where multiple-scattering effects are taken into account. These effects lead to a



14.13 Scattering time ratio $\tau_t^{(0)}/\tau_s^{(0)}$ versus electron density N for a heterostructure with IS ($N_i = 1 \times 10^9$ cm^{-2}, $z_i = 0$ Å) and IRS ($\Delta = 3$ Å, $\Lambda = 30$ Å) as the solid line. The dashed line (IS) and the dashed-dotted line (IRS) represent a single scattering mechanism. The dotted line for the two scattering mechanisms is calculated within the random-phase approximation with $G(q) = 0$.



11.14 Resistance ratio $\rho_0(B = B_c)/\rho_0(B = 0)$ versus electron density N for a heterostructure with IS ($N_i = 1 \times 10^9$ cm^{-2}, $z_i = 0$ Å) and with IRS ($\Delta = 3$ Å, $\Lambda = 30$ Å) as the solid line. The dashed line (IS) and the dashed-dotted line (IRS) represent a single scattering mechanism.

MIT at a critical density N_c, in this case at $N_c \approx 5 \times 10^9$ cm^{-2}. Deviations between μ and μ_0 on the logarithmic scale used in Fig. 14.12 are visible for $N_c < N < 3 \times 10^{10}$ cm$^{-2} \approx 6N_c$.

In the following we study in more detail the peak mobility and the critical density for HSs with very high mobility. A similar analysis has been made for silicon MOSFET structures.[3,34] For the same IRS parameters as used in Fig. 14.12 we show in Fig. 14.15 the peak mobility μ_m versus peak electron density N_m. The impurity density is varied in order to change the peak mobility. The general behaviour is the same as for silicon MOSFET.[3] A recent experimental result[8] is also shown.

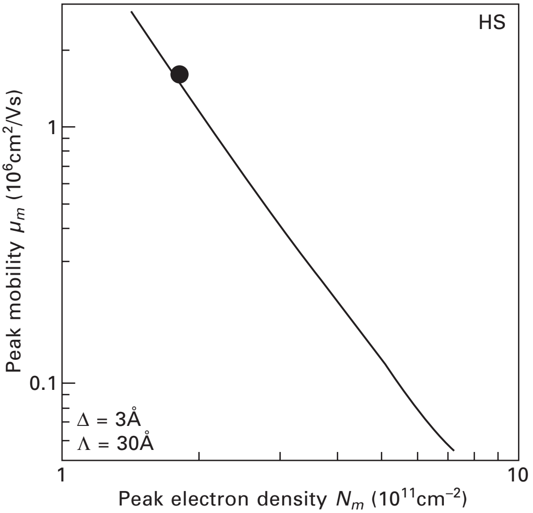
In Fig. 14.16 we show the dependence of the peak electron density N_m and of the critical density N_c as functions of the impurity concentration N_i. N_m and N_c increase with increasing N_i. Note that for $N_i \approx 1 \times 10^9$ cm^{-2} one impurity can localize 5 electrons.

14.5 Comparison with experimental results

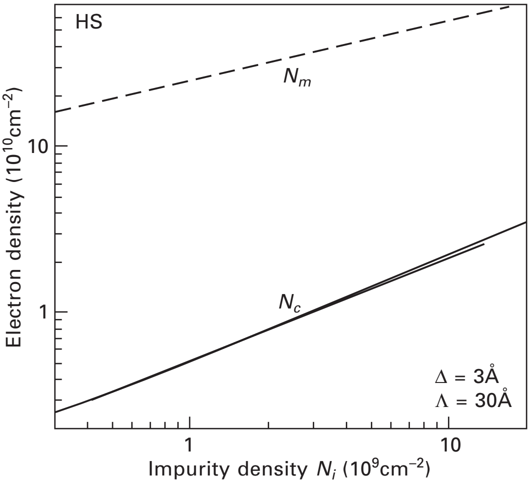
In this section we discuss experimental results for Si/SiGe QWs and HSs in connection with the transport theory described above. The discussion is sometimes quantitative, sometimes qualitative. A more detailed comparison with particularly interesting experiments will be published elsewhere.

14.5.1 Experiments with quantum wells

In Si QWs the IRS is reduced by a factor $(m^*/m_z)^2 \approx 0.043$ and one needs a small QW width in order to establish IRS in Si/SiGe QWs compared to



14.15 Peak mobility μ_m versus peak electron density N_m for a heterostructure with IS (N_i is varied, z_i = 0 Å) and with IRS (Δ = 3 Å, Λ = 30 Å). The solid dot represents the experimental result of Ref. 8.



14.16 Electron densities N_m and N_c versus impurity density N_i of heterostructures with IRS (Δ = 3 Å, Λ = 30 Å) and IS with impurities at the interface (z_i = 0 Å).

GaAs QWs.[5] It seems that there exists only one transport study where IRS in QWs has clearly been seen in experiment.[53] The mobility for QWs with well width 13 Å < L < 100 Å and with $N \approx 2 \times 10^{11}$ cm^{-2} was measured and showed a dramatic change near $L \approx L_c \cong 30$ Å. This is in good agreement with our prediction; compare Fig. 14.9 with Fig. 5 in Ref. 53.[54] The mobility was measured for temperatures higher than 20 K and a clear conclusion concerning a MIT cannot be drawn at such high temperatures. Variable-range hopping was seen for a QW of width $L \approx 13$ Å,[53] which indicates an insulating behaviour. From the linear temperature dependence of the mobility for a sample with $L \approx 40$ Å it was concluded that $\Delta \approx 7.9$ Å and $\Lambda \approx 129$ Å. This would imply $k_F \Lambda \approx 1.02$, which is intermediate between short-range and long-range behaviour. However, we are not completely convinced that the temperature dependence of the mobility can give accurate values of Λ, especially because a factor of 2 is missing in the numerical approach,[26] which was used in Ref. 53, and because the effect of the LFC has been neglected. The inclusion of the LFC leads to a lower mobility and would certainly reduce the value of Δ to explain the experimental data. Unfortunately, the single-particle relaxation time was not measured. Another method to determine the parameters of IRS for QWs of different width was proposed recently.[55] But this method only works when IRS is present and IS can be ignored.

In principle the measurement of the linear temperature dependence of the mobility is a very interesting way to show that IRS is present in remote-doped samples. We remember that remote doping does not contribute to the linear temperature dependence of the mobility. In fact, we stress that there are not many measurements of the linear temperature dependence of the mobility in Si/SiGe, probably because of the dominance of IS in remote-doped samples, where the linear temperature dependence is exponentially suppressed.[29]

The linear temperature dependence of the mobility has been studied in a Si/SiGe QW of width $L = 150$ Å and a spacer $z_i = -140$ Å.[56] A very weak temperature dependence of the resistance has been found, due to remote doping with $\tau_t/\tau_s \approx 15$, much weaker than in MOSFET structures, which is in agreement with the theoretical predictions.[29] We believe that some impurities at the interface are responsible for the linear temperature dependence.

The mobility of thin QWs with $L = 50$ Å has been studied[57] but the mobility $\mu \approx 10^4$ cm^2/Vs of these structures was too low to attribute this mobility to IRS, which would be at least $\mu \approx 10^5$ cm^2/Vs for the density used in the experiment – see Fig. 14.4. Of course, the argument also depends on the value of Λ. We conclude that in the samples of Ref. 57 IS was still the most important scattering mechanism.

An increase of the mobility with increasing electron density was measured in remote-doped samples.[58] This increase is in agreement with theory[5] – see Fig. 14.2. Remote doping effects have been studied in Ref. 59 and, as expected,[5] it was found that the mobility increases strongly with the spacer width – see Fig. 14.2. In fact, we find that the analytical formula[19] $\tau_t^{(0)} \propto (k_F\,|z_i|)^3$ describes reasonably well the experimental results of Ref. 59: an increase of the mobility by a factor of 5.5 and a decrease of the density by a factor of 1.7 by increasing the spacer from 170 Å to 400 Å. That work[59] showed that remote doping was the essential scattering mechanism in the samples studied. High mobility values in the range $\mu \approx 4 \times 10^5$ cm^2/Vs have been reported[60] for QW width smaller than a critical thickness L_d, where no dislocation could enter the silicon channel. This critical thickness depends on the x content in the Si_xGe_{1-x}.

In Ref. 61 it was argued, using the measured mobility of a QW versus electron density, that a screening breakdown at low electron density gives rise to a reduced mobility, if compared with theory.[5] A MIT was *extrapolated* (but not measured) to happen at $N_c \approx 7 \times 10^{10}$ cm^{-2} and it was argued that the density of states should approach zero at the density of the MIT. While impurity bands might exist in Si/SiGe, we think that in the samples of Ref. 61 impurity bands are not present due to the remote doping and the small binding energy for such impurities. We stress that it was argued in detail that for remote doping the disorder-induced density of states effects cancel out in the mobility, when calculated in lowest order.[38] This means that the physical picture used in Ref. 61 cannot explain the reduced mobility found in experiment. In addition, it was shown later[62] by measurements with a similar sample and by comparison with theory[5] that the reduced mobility found in Ref. 61 can be explained by many-body effects via a finite LFC. This is in agreement with our transport theory.[34] The inclusion of many-body effects leads to a stronger density variation of the mobility in low-density samples – see Fig. 14.1 for QWs and Fig. 14.11 for HSs. The estimated MIT at $N_c \approx 7 \times 10^{10}$ cm^{-2} is in perfect agreement with the mode-coupling theory, where the density of states is constant. The detailed comparison between experiment and the theory is published elsewhere.[63]

Spin polarization effects in a parallel magnetic field in Si/SiGe QWs have been studied for QW widths $L = 150$ Å,[62] and for $L = 200$ Å.[64] The QWs were remote doped. Weak orbital effects in the magnetoresistance are seen in these structures. The resistance ratio $\rho(B_c)/\rho(B = 0) \approx 1.3$, found in both QWs, is in excellent agreement with the theoretical prediction[25] for remote doping – see Fig. 14.7.

A MIT transition at high electron density $N_c \approx 4 \times 10^{11}$ cm^{-2} was found in a low-mobility Si/SiGe structure, which was remote doped, corresponding to a long-range random potential.[65] In the metallic regime a linear temperature dependence of the mobility has been found, which means that impurities *must* be present in the QW.[29] This is in agreement with the arguments of the authors of Ref. 65 that the transport properties in the studied sample were not explainable by long-range disorder alone. We believe that both nearby *and* remote impurities are present in the sample used in Ref. 65.

The MIT, the temperature dependence of mobility, and the magnetoresistance in a parallel magnetic field were also studied in a series of publications using essentially three similar samples.[66–68] The MIT in a high-mobility sample occurred around $N_c \approx 3 \times 10^{10}$ cm^{-2}.[66] In Table 14.1 we show characteristic parameters of the structure used. The remote-doped sample showed a resistance ratio $\rho(B_c)/\rho(B = 0) \approx 1.9$.[66,68] This is in good agreement with our results for remote doping, if we assume some additional IS from a small number of impurities in the well or at the interface – see Fig. 14.7. Scaling behaviour[24] of the magnetoresistance $\rho(B)$ in a parallel magnetic field was also verified in experiment – see Fig. 3b in Ref. 68. Most interesting, it was shown that $\rho(B_c)/\rho(B = 0) >> 1$ for $N \to N_c$.[68] In the framework of the

Table 14.1 Experimental values for the characteristic mobility (at a given density) and the critical density N_c of the metal–insulator transition found in Si/SiGe QWs and HSs

Type	Mobility (at width/density)	$N_c/10^{11}$cm^{-2}	Year	Ref.
QW	3×10^4 cm^2/Vs (at $L = 100$ Å)	2.0	1996	[53]
QW	6×10^4 cm^2/Vs (at 6.2×10^{11} cm^{-2})	4.0	2003	[65]
QW	1.9×10^5 cm^2/Vs (at 1.5×10^{11} cm^{-2})	0.32	2005	[66]
QW	1.9×10^5 cm^2/Vs (at 1.5×10^{11} cm^{-2})	0.32	2007	[67]
QW	1.9×10^5 cm^2/Vs (at 1.3×10^{11} cm^{-2})	0.28	2008	[68]
QW	2.8×10^4 cm^2/Vs (at 1.3×10^{11} cm^{-2})	0.95	2008	[68]
HS	6.3×10^4 cm^2/Vs (at 4.1×10^{11} cm^{-2})	2.35	2004	[73]
HS	5.5×10^4 cm^2/Vs (at 4.2×10^{11} cm^{-2})	1.90	2007	[74]
HS	1.6×10^6 cm^2/Vs (at 3.0×10^{11} cm^{-2})	<0.25	2009	[8]

magnetotransport theory including the MIT[39], this means that $N_c(g_s = 1) > N_c(g_s = 2)$ – see, for instance, Fig. 14.10 for IRS. For $N_c(g_v = 1) < N_c(g_s = 2)$ one expects $\rho(B_c)/\rho(B = 0) \rightarrow 0$ at $N_c(g_s = 2)$. A divergence of the effective mass was not observed in Ref. 68, which means that the physics of the MIT observed in Ref. 68 must be a disorder-induced MIT. What is also in favour of a disorder-induced MIT are the different values of N_c found in experiment[65–68] – see Table 14.1. Reference 68 represents certainly the most important paper concerning the in-plane magnetic field effect in Si/SiGe samples.

For the linear temperature dependence a negative value for $C(\alpha, N) < 0$ has been found, though for higher temperatures $T \approx \varepsilon_F/k_B$.[67] At those temperatures the expansion of Eq. 14.11 is no longer valid.

14.5.2 Experiments with heterostructures

Atomic-force microscopy was used to determine the IRS parameters for HSs.[69] Different roughness scales have been identified with this method: growth roughness with $\Lambda = 1000$ Å and atomic scale roughness with $\Lambda = 10$ Å to 100 Å. However, IRS with $\Lambda < 50$ Å cannot be seen with atomic-force microscopy. Therefore the real length scale of the atomic-scale roughness was not determined. Comparing mobility calculations with mobility data, it was concluded[69] that IRS is not relevant for the mobility and that the mobility was determined by IS and not by IRS. The fact that $\Lambda = 1000$ Å does not show up in the mobility is due to the rapid decay of the random potential with wave number, possibly exponentially as in Eq.(14.4).

The transport scattering time and the single-particle scattering time have been determined and a ratio $\tau_t/\tau_s \approx 14$–26 was reported for four samples with a high mobility around $\mu \approx 10^5$ cm^2/Vs.[70] Using the formula[19] for remote doping $\tau_t/\tau_s = (2k_F z_i)^2 = 4\pi N z_i^2$ we can use the numbers given for τ_t/τ_s and N in order to obtain z_i. From the experimental numbers $\tau_t/\tau_s = 26$ and $N = 7 \times 10^{11}$ cm^{-2} we get $z_i = 172$ Å, which is in reasonable agreement with the spacer width of 100 Å, a doping width of about 100 Å and an extension of the electron gas of 50 Å: $z_{i,\text{eff}} = (100 + 40 + 25)$ Å. But of course, the variation[70] $\tau_t/\tau_s \approx 14$–26 indicates that for some samples there might be IS from impurities near the interface, reducing τ_t/τ_s. A measurement of the density dependence could give additional information. But the essential scattering mechanism in these samples is certainly remote IS – see Fig. 14.5. We find it astonishing that the elementary analysis using the analytical formula $\tau_t/\tau_s = 4\pi N z_i^2$ was not made in Ref. 70.

The ratios $\tau_t/\tau_s \approx 2.4$–$2.9$ for $N \approx 9 \times 10^{11}$ cm^{-2} found[71] for intermediate remote doping $z_i = 40$ Å show that one is intermediate between a short-range and a long-range random potential – see Fig. 14.5.

Ratios $\tau_t/\tau_s \approx 18$ have been reported for samples with $\mu \approx 3.5 \times 10^5$

cm^2/Vs for $N = 5 \times 10^{11}$ cm^{-2}.[72] Using the analytical formula, we find z_i = 169 Å, in reasonable agreement with the doping profile where a spacer of nominal width 150 Å was used. We conclude that the mobility is essentially determined by remote IS.

The modification of the electron density with a gate was used to study the mobility and to observe a MIT at $N_c = 2.35 \times 10^{11}$ cm^{-2}.[73] The variation of the density was limited in this sample to 2.13×10^{11} $cm^{-2} < N < 4.24 \times 10^{11}$ cm^{-2}, and the sample was remote doped. In the metallic range 4.06×10^{11} $cm^{-2} < N < 4.24 \times 10^{11}$ cm^{-2} a linear temperature dependence of the mobility was found, in agreement with the screening theory of an interacting electron gas.[26,27] We find that the strength of this linear temperature dependence is reduced compared to silicon MOSFET structures, which is a sign of some remote doping.[29] While the sample used in Ref. 73 was remote doped, we conclude that a part of the resistance must be due to IS from impurities located at the interface of the HS. We stress that we deduce this from the existence of the linear temperature dependence found in the experiment. However, nothing about the position of the impurities concerning this linear temperature dependence was said in the original paper. We also note that the mobility of the sample used in Ref. 73 was considerably reduced – see Table 14.1, compared to the samples used in Ref. 68. This is a hint that impurities at the interface are present.

An undoped HS was studied in Ref. 74 where the electron gas was capacitively induced and a MIT was observed at $N_c \approx 1.9 \times 10^{11}$ cm^{-2}. We believe that the disorder in this structure is due to charged impurities located in the Al_2O_3 oxide, which has a distance of α = 490 Å from the two-dimensional electron gas. This corresponds to a long-range random potential. Due to the long-range nature of the disorder, the mobility scale at the MIT is strongly enhanced, $\sigma(N \geq N_c) \approx \sigma_0(N/N_c - 1)$ with $\sigma_0 >> e^2/h$, as predicted in Ref. 38. This is, in fact, seen in the experiment,[74] where $\sigma_0 \approx 5e^2/h$ was measured. The sample had a low mobility $\mu \leq 5.5 \times 10^4$ cm^2/Vs – see Table 14.1 – which indicates that the impurity density in the Al_2O_3 is high. We conclude that there is a large potential for mobility increase in the future.[75]

Given the recent high mobility sample with $\mu \approx 1.6 \times 10^6$ cm^2/Vs,[8] it seems that the gating process at low temperatures is now under much better control. It is astonishing that Shubnikov–de Haas studies have not been published for this sample, but we expect that the relation $\tau_t/\tau_s >> 1$ holds.[76]

14.5.3 Discussion of the metal–insulator transition

IRS is very important in thin QWs and we believe that an IRS-induced MIT was seen in these thin silicon QWs.[53] Unfortunately, the result of this paper has not yet been confirmed by additional measurements on similar samples and at low temperatures. In all the other Si/SiGe samples showing a MIT and discussed in Table 14.1 we believe that IS is the essential scattering mechanism. In Table 14.1 we have shown some characteristics of Si/SiGe samples showing a MIT. More results will be presented elsewhere.[54,63]

In general one can say that the mobility at a given and fixed electron density increases with decreasing N_c and, equivalently, with decreasing N_i. But it is also clear from the data shown in Table 14.1 that one has to know the position of impurities for a detailed comparison with theory. We have concluded that in some samples there are remote impurities *and* impurities in the QW (or at the interface). This clearly shows that more systematic work is necessary, including the study of the single-particle relaxation time. It is clear from Table 14.1 that a disorder-induced MIT exists. Nevertheless, there are still many people who believe that two-dimensional electrons are always localized, at least at zero temperature, due to weak localization effects. This belief exists even in face of record high mobilities in GaAs structures[2] and Si/SiGe structures.[8] This is the reason why some people speak about an 'apparent MIT'.[73]

We believe that the claim[32,46] of an interaction-induced MIT at $N_{c,int}^{Si(100)} \approx 0.8 \times 10^{11}$ cm^{-2} in silicon MOSFET structures needs confirmation by other groups. If we take seriously the arguments given in Ref. 32, we are led to the conclusion that for silicon MOSFET structures and for stronger disorder, a disorder-induced MIT exists with $N_c > N_{c,int}^{Si(100)}$.[1,34] If disorder becomes smaller and smaller, the MIT becomes finally interaction induced with $N_c = N_{c,int}^{Si(100)}$. However, we stress that in the (111) silicon MOSFET structure studied in Ref. 46 disorder effects are very strong, as was shown in Ref. 41. Without disorder no vanishing mobility would exist.

We believe that in the Si/SiGe system (see Table 14.1) a disorder-induced MIT is found with N_c depending on disorder. An interaction-induced MIT should always occur at the same electron density. Taking into account the change of the background dielectric constant, which results in a factor of 7.7/12.5, we would expect the interaction-induced MIT in Si/SiGe to be $N_{c,int}^{Si/SiGe} \approx 3 \times 10^{10}$ cm^{-2}. While this number is in agreement with some results obtained in Ref. 66, no mass divergence was observed. In addition, in the high-mobility sample[8] no mobility anomaly was observed at this density, and a high mobility $\mu = 5.6 \times 10^4$ cm^2/Vs was reported at $N = 2.5 \times 10^{10}$ cm^{-2}, a density below $N_{c,int}^{Si/SiGe}$. We expect that the future study of high-mobility Si/SiGe structures will be very interesting, especially with respect to the MIT. We claim that the long-range nature of the disorder leads to a MIT as described long ago in Ref. 38.[75]

14.6 Discussion and future trends

In the literature often a 'percolation type of conduction' is used in connection with transport near the MIT in two-dimensional systems. Whenever the mobility depends on the spin degree of freedom, which can be tested using a magnetic field, quantum effects are necessary to explain these effects. As has been seen in the present chapter, the mobility depends strongly on the magnetic field. This rules out all percolation theories. From this we conclude that the percolation picture is irrelevant. It is strange that such a simple argument is not used in the community.

The theory for IRS[5] in QWs predicted non-monotonic mobility as a function of the density – see Figs 14.3 and 14.4. IRS as dominant scattering mechanism in thin QWs[13] is well established in GaAs QWs[14,15] and InAs QWs.[16] The importance of IRS in thin Si/SiGe QWs was shown in Ref. 53. However, the non-monotonic mobility of IRS has not yet been seen in any experiment. Only an increase of mobility with density has been found.[14–16] The minimum of the mobility would give information about Λ, (see Fig. 14.4). Only recently was a decrease of the mobility with increasing density seen in an experiment, though with an InAs/GaSb superlattice.[77] However, the subsequent expected increase of the mobility with increasing density was not established in the experiment. It remains to be seen if this non-monotonic mobility can be found in experiment, for instance in Si/SiGe. The electron density variation should be sufficiently large to be able to study this phenomenon.

There are indications[53] of a MIT due to IRS but much more work is necessary. It seems that a MIT driven by IRS was seen in InGaAs/InP superlattices[78] and in AlAs QWs.[79] We believe that these MITs need confirmation. IRS is important in thin QWs and study of the IRS parameters gives information about interfaces. These parameters are also very important for quantum well lasers.[80] There IRS in different subbands must be considered[81] and eventually finite confinement effects might be important.[82]

Concerning the linear temperature dependence, we would like to point out that if the LFC is very large $G(q = 2k_F) > 1$ and $[1 - G(q \approx 2k_F)]F(2k_F) + 2k_F/q_s > 0$ one finds $C(\alpha, N) < 0$.[30] Of course, such a coefficient cannot be found within the RPA or using a LFC of Hubbard type. It cannot be excluded that such a behaviour will be observed in the future in systems with low densities. Until now it seems that $C(\alpha, N) < 0$ has not yet been seen in an experiment.

We predict that in the near future we shall see gated and undoped Si/SiGe HSs having a mobility at low temperatures of up to $\mu \approx 10^7$ cm^2/Vs. It was argued that screening in Si/SiGe is much better[40] than in GaAs due to the larger valley degeneracy in silicon. The MIT in such high-mobility samples will shift to very low densities. One needs to characterize these samples carefully in order to obtain information about the disorder of such undoped samples, where the origin of disorder is not yet known. We expect that in these HSs at low electron density impurities at the interface are important and at higher density IRS and alloy scattering will become important. These scattering mechanisms are associated with a linear temperature dependence. This effect should be used to get microscopic insight and will give new importance to the linear temperature dependence of the mobility.

Eventually new ground states will be found at very low electron densities. Nevertheless, one should understand better the disorder-induced MIT. We stress that the disorder-induced MIT is a quite general phenomenon, recently also found in the two-dimensional electron gas at the interface of ZnO/MgZnO HSs.[83–85] In Table 14.2 we give some characteristic transport data of such structures. One clearly sees that, due to the large critical density around $N_c \approx 5 \times 10^{11}$ cm^{-2},[83,84] there is plenty of room for a large mobility increase by reducing the disorder in order to get $N_c \approx 1 \times 10^{11}$ cm^{-2}, as already found recently.[85] With an effective mass of $m^* \approx 0.29\, m_e$ and $\varepsilon_L \approx 8.5$, the interaction-induced MIT in this system should be at $N_{c,int}^{ZnO/MgZnO} \approx 1.53 \times 10^{11}$ cm^{-2}. It is evident from this number and the numbers given in Table 14.2 that the MIT in these samples is induced by disorder and not by interaction.[83–85] We expect much new and exciting data in the future from these oxides. Recently it was shown that the electron gas in ZnO/MgZnO HSs with a MIT can be described by IS and IRS.[86]

A large volume of data on disorder-induced MIT exists for GaAs/AlGaAs HSs, which occurs at very low electron density due to the high-mobility samples used in experiments. We are not going into this topic here, but in Ref. 38 the transport properties including a MIT were described for such structures.

Transport properties of the two-dimensional electron gas at a silicon (111) surface with high valley degeneracy, $g_v = 6$, will become very important in the future. Screening is much better there due to the larger valley degeneracy, and the mobility should be strongly enhanced.[23,40,42,87] Anisotropic electron gases could also be a topic in the future, for instance for the electron gas at the silicon (110) surface. There, electron transport is anisotropic, mobility ratios of 10 can be achieved,[88] and this might become a technological issue. We suggest the study of Si(110)/Si(110)Ge structures and Si(111)/Si(111) Ge structures.

Table 14.2 Experimental values for the characteristic mobility (at a given electron density) and the critical density N_c of the metal–insulator transition found in ZnO/MgZnO HSs

Type	Mobility (at density)	$N_c/10^{11}$cm^{-2}	Year	Ref.
HS	2.5×10^3 cm^2/Vs (at 10×10^{11} cm^{-2})	4.0	2007	[83]
HS	5.0×10^3 cm^2/Vs (at 12×10^{11} cm^{-2})	5.0	2008	[84]
HS	2.2×10^4 cm^2/Vs (at 5×10^{11} cm^{-2})	1.1	2010	[85]
HS	2.0×10^4 cm^2/Vs (at 5×10^{11} cm^{-2})	0.9	2010	[85]

14.7 Conclusions

More *systematic* measurement should be done with thin Si/SiGe QWs. The first aim should be to determine interface-roughness parameters. Second, thin quantum wells should be studied in order to observe a metal–insulator transition due to interface-roughness scattering. Studies of Dingle temperature should be made together with transport measurements. Symmetrically remote doped quantum wells should be used.

For the mobility of HS a large density range should be studied in order to obtain information about the surface-roughness parameters. Low density HSs are most important for the MIT. The insulator-gate field-effect transistor is most promising.[8]

The authors of Ref. 32 reduce the discussion of the metal–insulator transition to an interaction-driven one, described by a mass divergence. This makes no sense in the light of most experimental data on the MIT in two-dimensional systems available in the literature. In our opinion mass divergence is not the origin of the MIT (the vanishing of the conductivity at zero temperature) in silicon MOSFETs. We believe that if mass divergence exists, then a disorder-induced MIT will occur nearby.

While great progress has been made during the last 25 years in increasing the mobility of the electron gas in Si/SiGe structures, the understanding of interface-roughness parameters has not really improved. Some theoretical results have been obtained and applied favourably to GaAs quantum wells. A lot needs to be done for Si/SiGe structures.

The final result of such efforts would be, on the theoretical side, a trustable transport theory going beyond lowest order in disorder and applicable to dilute systems with $r_s > 1$, where many-body effects are strong. Concerning the experimental side we believe that systematic studies would allow one to increase the mobility considerably.[8,75] Understanding of the disorder would probably help researchers to observe and to understand more complex ground states. Also magnetic systems in connection with semiconductors could be considered.

There are still people who believe that two-dimensional electron systems are isolating at zero temperature due to weak localization effects. Therefore we want to present some numbers. The mobility of Si/SiGe structures has increased by a factor of 700 during the last 25 years. For the GaAs/AlGaAs system the mobility increased[2] by a factor of 60 in 21 years, from $\mu \approx 5 \times 10^5$ cm^2/Vs in 1983, for the sample used to study the fractional quantized Hall effect,[89] to $\mu \approx 31 \times 10^6$ cm^2/Vs in 2004.[90] In addition, within these 25 years the temperatures used to characterize the electronic properties of QWs and HSs were reduced from 1 K to about 10 mK, which means an additional factor of 100, this time in temperature. All these factors (and the experimental results related to these factors) are not in favour of the existence of an isolating ground state at zero temperature due to weak localization effects.[35] The facts are in favour of a MIT at a critical density N_c.[34] This evolution was already predicted some time ago in connection with the MIT in GaAs/AlGaAs.[38] We believe that many people have not understood that a *one-parameter* scaling *hypothesis* is not a one-parameter scaling *theory* and is not a one-parameter scaling *reality*. In experiments there exists a metallic phase in two dimensions. Sometimes, looking at experiments and extrapolating to zero temperature, it is more the question of whether there really exists an insulator phase. We suggest that studying the density dependence of the mobility is an efficient way to identify the metal–insulator transition.

It is quite obvious that technological breakthroughs,[2,8,90] in order to reduce disorder, are responsible for the mobility increase during the last years and for the discovery of new ground states. We believe that the study of disorder-dominated metallic and localized phases in two-dimensional electron systems with $r_s > 1$ will be a major topic in the coming years.

14.8 Acknowledgements

My post-doctoral studies have been financed by the Ernst von Siemens Stipendium and the Deutsche Forschungsgemeinschaft. I gratefully wish to thank Professor F. Koch in Munich for his generous support after my thesis work and for many fruitful discussions and helpful advice. I thank Professors J. B. Renucci and M. A. Renucci in Toulouse for advice over the last 19 years. It is also a pleasure to thank Professor V. T. Dolgopolov in Chernogolovka for discussions over the last 25 years.

14.9 References

1 Ando T., Fowler A. B. and Stern F., Electronic properties of two-dimensional systems. *Review of Modern Physics* 54 (2), 437–672 (1982).

2 Willett R. L., Pfeiffer L. N. and West K. W., Simple-layered high mobility field effect heterostructured two-dimensional electron device. *Applied Physics Letters* 89 (24), 242107 (2006).

3 Gold A., Peak mobility of silicon metal-oxide semiconductor systems. *Physical Review Letters* 54 (10), 1079–1082 (1985).

4 Abstreiter G., Brugger H., Wolf T., Jorke H. and Herzog H. J., Strain induced two-dimensional electron gas in selectively doped Si/Si_xGe_{1-x} superlattices. *Physical Review Letters* 54 (22), 2441–2444 (1985).

5 Gold A., Electronic transport properties of a two-dimensional electron gas in a silicon quantum well at low temperature. *Physical Review B* 35 (2), 723–733 (1987).

6 Schäffler F., Többen D., Herzog H.-J., Abstreiter G. and Holländer B., High-electron mobility Si/SiGe heterostructures: influence of the relaxed SiGe buffer layer. *Semiconductor Science and Technology* 7 (2), 260–266 (1992).

7 Schäffler F., High mobility Si and Ge structures. *Semiconductor Science and Technology* 12 (12), 1515–1550 (1997).

8 Lu T. M., Tsui D. C., Lee C.-H. and Liu C. W., Observation of two-dimensional electron gas in a Si quantum well with mobility of 1.6×10^6 cm^2/Vs. *Applied Physics Letters* 94 (18), 182102 (2009).

9 Pines D. and Nozières P., *The Theory of Quantum Liquids*. Benjamin, New York (1966).

10 Singwi K. S. and Tosi M. P., Correlations in electron liquids. *Solid State Physics* 36, 177–266 (1981).

11 Gold A., Local-field correction for the electron gas: effects of the valley degeneracy. *Physical Review B* 50 (7), 4297–4305 (1994).

12 Gold A., The local-field correction for the interacting electron gas: many-body effects for unpolarized and polarized electrons. *Zeitschrift für Physik B – Condensed Matter* 103 (3–4), 491–500 (1997).

13 Gold A., Metal-insulator transition due to surface-roughness scattering in a quantum well. *Solid State Communication* 60 (6), 531–534 (1986).

14 Sakaki H., Noda T., Hirakawa K., Tanaka M. and Matsusue T., Interface-roughness scattering in GaAs/AlAs quantum wells. *Applied Physics Letters* 51 (14), 1934–1937 (1987).

15 Gottinger R., Gold A., Abstreiter G., Weimann G. and Schlapp W., Interface-roughness scattering and electron mobilities in thin GaAs quantum wells. *Europhysics Letters* 6 (2), 183–188 (1988).

16 Bolognesi C. R., Kroemer H. and English J. H., Interface roughness scattering in InAs/AlSb quantum wells. *Applied Physics Letters* 61 (2), 213–216 (1992).

17 Prange R. E. and Nee T. W., Quantum spectroscopy of the low-field oscillations in the surface impedance. *Physical Review* 168 (3), 779 (1968).

18 Ando T., Screening effect and quantum transport in a silicon inversion layer in strong magnetic fields. *Journal of the Physical Society of Japan* 43 (5), 1616–1626 (1977).

19 Gold A., Scattering time and single-particle relaxation time in a disordered two-dimensional electron gas. *Physical Review B* 38 (15), 10798–10811 (1988).

20 Das Sarma S. and Stern F., Single-particle relaxation time versus scattering time in an impure electron gas. *Physical Review B* 32 (12), 8442–8444 (1985).

21 Okamoto T., Hosoya K., Kawaji S. and Yagi A., Spin degree of freedom in a two-dimensional electron liquid. *Physical Review Letters* 82 (22), 3875–3878 (1999).

22 Das Sarma S. and Wang E. H., Parallel magnetic field induced giant magnetoresistance in low density quasi-two-dimensional layers. *Physical Review Letters* 84 (24), 5596–5599 (1999).

23 Dolgopolov V. T. and Gold A., Magnetoresistance of a two-dimensional electron gas in a parallel magnetic field. *JETP Letters* 71 (1), 27–30 (2000).

24 Gold A. and Dolgopolov V. T., Resistivity of the spin-polarized two-dimensional electron gas: scaling and width effect. *Physica E* 17, 280–283 (2003).

25 Gold A., Single-particle relaxation time of the two-dimensional spin-polarized electron gas: remote doping. *Physica E* 17, 305–306 (2003).

26 Stern F., Calculated temperature dependence of mobility in silicon inversion layers. *Physical Review Letters* 44 (22), 1469–1472 (1980).

27 Gold A. and Dolgopolov V. T., The temperature of the maximal conductivity in Si MOS systems. *Journal of Physics C: Solid State Physics* 18 (16), L463–L466 (1985).

28 Gold A. and Dolgopolov V. T., Temperature dependence of the conductivity for the two-dimensional electron gas: analytical results for low temperatures. *Physical Review B* 33 (2), 1076–1084 (1986).

29 Gold A., Temperature dependence of mobility in $Al_xGa_{1-x}As$/GaAs heterostructures for impurity scattering. *Physical Review B* 41 (12), 8537–8540 (1990).

30 Gold A., Linear temperature dependence of the mobility in two-dimensional electron gases: many-body and spin-polarization effects. *Journal of Physics: Condensed Matter* 15 (2), 217–223 (2003).

31 Zala G., Narozhny B. N. and Aleiner I. L., Interaction corrections at intermediate temperatures: longitudinal conductivity and kinetic equation. *Physical Review B* 64 (21), 214204 (2001).

32 Kravchenko S. V. and Sarachik M. P., Metal–insulator transition in two-dimensional systems. *Reports on Progress in Physics* 67 (1), 1–44 (2004).

33 Gold A. and Götze W., The conductivity of strongly disordered two-dimensional systems. *Journal of Physics C: Solid State Physics* 14 (28), 4049–4066 (1981).

34 Gold A. and Götze W., Localization and screening anomalies in two-dimensional systems. *Physical Review B* 33 (4), 2495–2511 (1986).

35 Lee P. A. and Ramakrishnan T. V., Disordered electronic systems. *Review of Modern Physics* 57 (2), 287–337 (1986).

36 Rahimi M., Anissimova S., Sakr M. R., Kravchenko S. V. and Klapwijk T. M., *Physical Review Letters* 91 (11), 116402 (2003).

37 Belitz D., Gold A. and Götze W., Self-consistent current relaxation theory for the electron localization problem. *Zeitschrift für Physik B – Condensed Matter* 44 (4), 273–277, (1981).

38 Gold A., Metal–insulator transition in $Al_xGa_{1-x}As$/GaAs heterostructures with large spacer width. *Physical Review B* 44 (16), 8818–8824 (1991).

39 Gold A., The conductivity of the spin-polarized two-dimensional electron gas: exchange correlation and strong disorder effects. *JETP Letters* 72 (5), 274–277 (2000).

40 Gold A. and Antonie O., Magnetoresistance of a silicon MOSFET on the (111) surface in a parallel magnetic field. *International Journal of Modern Physics B* 21 (8–9), 1529–1534 (2007) (*Proceedings of the 17th International Conference on High Magnetic Fields in Semiconductor Physics*, Würzburg, Germany, August 2006).

41 Gold A., Scattering times in the two-dimensional electron gas on silicon (111) with a density dependent effective mass. *Journal of Physics: Condensed Matter* 19 (50), 506214 (2007).

42 Gold A., Fabie L. and Dolgopolov V. T., Two-band model for transport properties of silicon (111) MOSFET structures with high mobility. *Physica E* 40 (5), 1351–1353 (2008).

43 Gold A., Transport and cyclotron resonance theory for GaAs–AlGaAs heterostructures. *Zeitschrift für Physik B – Condensed Matter* 63 (1), 1–8 (1986).

44 Gold A., Allen S. J., Wilson B. A. and Tsui D. C., Frequency-dependent conductivity of a strongly disordered two-dimensional electron gas. *Physical Review* B 25 (6), 3519–3528 (1982).

45 Gold A., Götze W., Mazuré C. and Koch F., The dynamical conductivity of a Na^+ doped interfacial charge layer on silicon. *Solid State Communication* 49 (11), 1085–1088 (1984).

46 Shashkin A. A., Kapustin A. A., Deviatov E. V., Dolgopolov V. T. and Kvon Z. D., Strongly enhanced effective mass in dilute electron systems: system independent origin. *Physical Review B* 76 (24), 241302 (2007).

47 Götze W., Leutheusser E. and Yip S., Diffusion and localization in the two-dimensional Lorentz model. *Physical Review A* 25 (1), 533–399 (1982).

48 Paul S. K. and Basu P. K., Electron mobility in a Si_xGe_{1-x} quantum well limited by alloy-disorder scattering. *Journal of Applied Physics* 70 (7), 3977–3979 (1991).

49 Laikhtman B. and Kiehl R. A., Theoretical hole mobility in a narrow Si/SiGe quantum well. *Physical Review B* 47 (16), 10515–10527 (1993).

50 Stern F. and Laux S. E., Charge transfer and low-temperature electron mobility in a strained Si layer in $Si_{1-x}Ge_x$. *Applied Physics Letters* 61 (9), 1110–1113 (1992).

51 Monroe D., Xie Y. H., Fitzgerald E. A., Silverman P. J. and Watson G. P., Comparison of mobility-limiting mechanisms in high-mobility $Si/Si_{1-x}Ge_x$ heterostructures. *Journal of Vacuum Science and Technology B* 11 (4), 1731–1737 (1993).

52 Quang D. N., Tuoc V. N., Huan T. D. and Phong P. N., Low-temperature mobility of holes in Si/SiGe *p*-channel heterostructures. *Physical Review B* 70 (19), 195336 (2004).

53 Yutani A. and Shiraki Y., Transport properties of *n*-channel Si/SiGe modulation-doped systems with varied channel thickness: effects of the interface roughness. *Semiconductor Science and Technology* 11 (7), 1009–1014 (1996).

54 Gold A., *in preparation*.

55 Gold A., Interface-roughness parameters in InAs quantum wells determined from mobility. *Journal of Applied Physics* 103 (4), 043718 (2008).

56 Brunthaler G., Prinz A., Pilwein G., Bauer G., Brunner K., Abstreiter G., Dietl T. and Pudalow V. M., Semiclassical origin of the 2D metallic state in high mobility Si-MOS and Si/SiGe structures. *Proceedings of the 25th International Conference of Semiconductors,* Osaka, Japan, September 2000, editors: Mura N. and Ando T., Springer Proceedings in Physics, Vol. 87, Berlin, pp. 785–786 (2001).

57 Schuberth G., Schäffler F., Besson M., Abstreiter G. and Gornik E., High electron mobility in modulation-doped Si/SiGe quantum well structures. *Applied Physics Letters* 59 (25), 3318–3320 (1991).

58 Nelson S. F., Ismail K., Jackson T. N., Nocera J. J. and Chu J. O., Systematics of electron mobility in Si/SiGe heterostructures. *Applied Physics Letters* 63 (6), 794–796 (1993).

59 Ismail K., Meyerson B. S. and Wang P. J., High electron mobility in modulation-doped Si/SiGe. *Applied Physics Letters* 58 (19), 2117–2120 (1991).

60 Ismail K., LeGoues F. K., Saenger K. L., Arafa M., Chu J. O., Mooney P. M. and Meyerson B. S., Identification of a mobility-limitation scattering mechanism in modulation doped Si/SiGe heterostructures. *Physical Review Letters* 73 (25), 3447–3451 (1994).

61 Wilamowski Z., Sandersfeld N., Jantsch W., Többen D. and Schäfffler F., Screening breakdown on the route towards the metal–insulator transition in modulation doped Si/SiGe quantum wells. *Physical Review Letters* 87 (2), 026401 (2001).

62 Dolgopolov V. T., Deviatov E. V., Shashkin A. A., Wieser U., Kunze U., Abstreiter G. and Brunner K., Remote-doping scattering and the local-field correction in the 2D electron system in a modulation-doped Si/SiGe quantum well. *Superlattices and Microstructures* 33 (5–6), 271–278 (2003).

63 Gold A., Mobility and metal–insulator transition of the two-dimensional electron gas in Si/SiGe quantum wells, *J. Appl. Phys.* 108 (6), 063710 (2010).

64 Okamoto T., Ooya M., Hosoya K. and Kawaji S., Spin polarization and metallic behaviour in a silicon two-dimensional electron gas. *Physical Review* B 69 (4), 041202 (2004).

65 Olshanetsky E. B., Renard V., Kvon Z. D., Portal J. C., Woods N. J., Zhang J. and Harris J. J., Conductivity of a two-dimensional electron gas in a Si/SiGe

heterostructure near the metal–insulator transition: role of the short- and long-range scattering potential. *Physical Review B* 68 (8), 085304 (2003).

66 Lai K., Pan W., Tsui D. C., Lyon S. A., Mühlberger M. and Schäffler F., Two-dimensional metal–insulator transition and in-plane magnetoresistance in a high-mobility strained Si quantum well. *Physical Review B* 72 (8), 081313(R) (2005).

67 Lai K., Pan W., Tsui D. C., Lyon S. A., Mühlberger M. and Schäffler F., Linear temperature dependence of the conductivity in Si two-dimensional electrons near the apparent metal-to-insulator transition. *Physical Review B* 75 (3), 033314 (2007).

68 Lu T. M., Sun L., Tsui D. C., Lyon S., Pan W., Mühlberger M., Schäffler F., Liu J. and Xie Y. H., In-plane field magnetoresistivity of Si two-dimensional electron gas in Si/SiGe quantum wells at 20 mK. *Physical Review B* 78 (23), 233309 (2008).

69 Feenstra R. M., Lutz M. A., Stern F., Ismail K., Mooney P. M., LeGoues F. K., Stanis C., Chu J. O. and Meyerson B. S., Roughness analysis of Si/SiGe heterostructures. *Journal of Vacuum Science and Technology B* 13 (4), 1608–1612 (1995).

70 Többen D., Schäffler F., Zrenner A. and Abstreiter G., Magnetotransport measurements and low-temperature scattering times of electron gases in high-quality $Si/Si_{1-x}Ge_x$ heterostructures. *Physical Review B* 46 (7), 4344–4347 (1992).

71 Stöger G., Brunthaler G., Bauer G., Ismail K., Meyerson B. S., Lutz J. and Kuchar F., Shubnikov–de Haas oscillations under hot-electron conditions in $Si/Si_{1-x}Ge_x$ heterostructures. *Physical Review B* 49 (15), 10417–10425 (1994).

72 Ismail K., Arafa M., Saenger K. L., Chu J. O. and Meyerson B. S., Extremely high electron mobility in Si/SiGe modulation doped heterostructures. *Applied Physics Letters* 66 (9), 1077–1079 (1995).

73 Lai K., Pan W., Tsui D. C. and Xie Y.–H., Observation of the apparent metal–insulator transition of high mobility electron systems in Si/SiGe heterostructures. *Applied Physics Letters* 84 (2), 302 (2004).

74 Lu T. M., Liu J., Kim J., Lai K., Tsui D. C. and Xie Y. H., Capacitively induced high mobility two-dimensional electron gas in undoped $Si/Si_{1-x}Ge_x$ heterostructures with atomic-layer-deposited dielectric. *Applied Physics Letters* 90 (18), 182114 (2007).

75 Gold A., Mobility of the non-polarized and the spin-polarized electron gas in Si/SiGe heterostructures: remote impurities. *Europhysics Letters* 92 (6), 67002 (2010).

76 Gold A., Metal–insulator transition in Si/SiGe heterostructures: mobility, spin polarization and Dingle temperature, *Semiconductor Science and Technology* 26 (3), in press (2011).

77 Szmulowicz F., Elhamri S., Haugan H. J., Brown G. J. and Mitchel W. C., Carrier mobility as a function of carrier density in type II InAs/GaSb superlattices. *Journal of Applied Physics* 105 (7), 074303 (2009).

78 Pusep Yu. A., Gozzo G. C. and LaPierre R. R., Interface roughness in short-period InGaAs/InP superlattices. *Applied Physics Letters* 93 (24), 242104 (2008).

79 Gold A., Mobility of thin AlAs quantum wells: theory compared to experiment. *Applied Physics Letters* 92 (8), 082111 (2008).

80 Lhuillier E., Ribet-Mohamed I., Rosencher E., Gilles Patriarche G., Buffaz A., Berger V. and Carras M., Interface roughness transport in terahertz quantum cascade detectors. *Applied Physics Letters* 96 (6), 061111 (2010).

81 Gold A., Theory of interface-roughness induced subband-edge energy renormalization in thin quantum wells. *Solid State Communication* 70 (3), 371–374 (1989).

82 Gold A., Barrier penetration effects for electrons in quantum wells: screening, mobility, and shallow impurity states. *Zeitschrift für Physik B – Condensed Matter* 74 (1), 53–65 (1989).

83 Shimotani H., Asanuma H., Tsukazaki A., Ohtomo A., Kawasaki M. and Iwasa Y., Insulator-to-metal transition in ZnO by electric double layer gating. *Applied Physics Letters* 91 (8), 082106 (2007).

84 Tsukazaki A., Ohtomo A., Chiba D., Ohno Y., Ohno H. and Kawasaki M., Low-temperature field-effect and magnetotransport properties in a ZnO based heterostructure with atomic-layer-deposited gate dielectrics. *Applied Physics Letters* 93 (24), 241905 (2008).

85 Nakano M., Tsukazaki A., Ohtomo A., Ueno K., Akasaka S., Yuji H., Nakahara K., Fukumura T. and Kawasaki M., Electric-field control of two-dimensional electrons in polymer-gated oxide semiconductor heterostructures. *Advanced Materials* 22 (8), 876–879 (2010).

86 Gold A., Transport properties of the electron gas in ZnO/MgZnO heterostructures. *Applied Physical Letters* 96 (24), 242111 (2010).

87 Gold A., Theory of transport properties of the sixfold degenerate two-dimensional electron gas at the H–Si (111) surface. *Physical Review B*, 82 (19), 195329 (2010).

88 Khrapai V. T. and Gold A., Transport properties of the two-dimensional interacting anisotropic electron gas at zero temperature. *Physical Review B* 81 (8), 085309 (2010).

89 Stormer H. L., Chang A., Tsui D. C., Wang J. C. M., Gossard A. C. and Wiegmann W., Fractional quantization of the Hall effect. *Physical Review Letters* 50 (24), 1953–1956 (1983).

90 Xia J. S., Pan W., Vicente C. L., Adams E. D., Sullivan N. S., Stormer H. L., Tsui D. C., Pfeiffer L. N., Baldwin K. W. and West K. W., Electron correlation in the second Landau level: a competition between many nearly degenerate quantum phases. *Physical Review Letters* 93 (17), 176809 (2004).

15

Transport properties of silicon–germanium (SiGe) nanostructures and applications in devices

F. SCHÄFFLER, Johannes Kepler Universität, Austria

Abstract: Mainstream electronics for the foreseeable future will be based on Si substrates and on transistors that rely on the manipulation of charged carriers. Hence, engineering of the transport properties by the introduction of Si-based heterostructures is, on the one hand, a strongly application-driven field. On the other hand, many electronic and spintronic material aspects, such as the influence of the valley degeneracy on the spin coherence times, can be addressed by transport experiments in low-dimensional Si/SiGe heterostructures. It is the purpose of this chapter to address carrier transport in SiGe heterostructures from its very basics to state-of-the-art devices.

Key words: SiGe heterostructures, modulation doping, mobility, two-dimensional carrier gas, quantum Hall effect, quantum point contacts, valley degeneracy.

15.1 Introduction

In 1985, strain in silicon-based heterostructures has been recognized as the single most important parameter for controlling the band alignments and the symmetries of the band structure.[1] Since then, strain engineering has become a powerful and versatile method for tailoring the structural, electronic, optical and spin properties in strained semiconductor heterosystems. In particular, carrier confinement in the SiGe heterosystem was soon recognized as a potential means for overcoming some of the physical limitations of this material system with its disadvantageous indirect band gap and rather mediocre carrier mobilities. Regarding the indirect gap, zone folding and low-dimensional carrier confinement schemes are being pursued in order to enhance radiative recombination, and to thus fulfill a precondition for efficient Si-based light emitters. Carrier mobilities, on the other hand, can be enhanced by influencing the scattering processes, and/or by modifying the effective masses. Again, carrier confinement and strain engineering are key issues that led to dramatic improvements of the carrier mobilities in Si-based heterostructures in recent years. In addition, spin properties have become an extremely active field of research, because of the prospects of quantum computing (QC) and quantum cryptography. In this respect, Si-based heterostructures are outstanding, because of the extremely long intrinsic spin coherence times in Si. They can be enhanced even more by depletion of the ^{29}Si isotope,[2] which is the only one with a nuclear spin. Still, many of the proposed QC schemes require the simultaneous manipulation of spins and charges in a highly integrated device environment that must sustain the extraordinary spin properties of Si.[3]

Any of the aforementioned research activities are of basic interest. There is, however, a strong application aspect to all Si-related research activities, which is based on the absolute market dominance of Si-based electronic devices. Any prospects of adding new functionalities to the ultra-high-integration capabilities of Si-based semiconductor technology are therefore of potential commercial impact. Thus, a main driving force for the development of Si-based optoelectronic or spintronic functionalities is their compatibility with standard Si technologies.

Still, mainstream Si electronics for the foreseeable future will be based on transistors that rely on the manipulation of charged carriers. Carrier transport is therefore of utmost relevance, and even comparably small improvements can become essential in a situation where both technological and material limitations jeopardize the straightforward extension of Moore's law[4] into its fifth decade. As a matter of fact, it was mainly mobility enhancement by strain engineering that saved Moore's law at around 2003, when Intel first utilized SiGe in its process flow.[5] Meanwhile, the introduction of high-*k* dielectrics and metal gates have solved for the time being two of the most severe problems on the international road map[6], but simultaneously strain engineering is advancing rapidly. Intel, for example, introduced in 2009 with its 32 nm technology already the fourth generation of its strained silicon process.[7,8]

This chapter is concentrated on transport in Si-based heterostructures. The first part gives an introduction to the properties of the carrier mobility as a function of the strain-dependent scattering mechanisms and variations of the effective mass. More detailed treatments of the influence of strain on the band structure, and thus on the effective masses, can be found elsewhere in this book. The second part gives an overview of the band alignment under strain, and its influence on the effective masses. The main emphasis is put on biaxially in-plain strained films on a (001) substrate. The third part deals with low-dimensional transport in high mobility n-type and p-type modulation-doped quantum well structures (MODQWs). Emphasis is put on n-MODQWs and the physical properties of the electrons in strained Si channels, which became accessible in the last few years due to the progress made regarding the carrier mobilities. The final section deals with the concepts and results of strain engineering in mainstream Si devices. For further information on transport in Si/SiGe heterostructures the reader is referred to Chapter 14 of this book, and to two review articles (Refs 9 and 10) and two data reviews (Refs 11 and 12).

15.2 Basic transport properties of strained silicon–germanium (SiGe) heterostructures

Within the framework of the classical Drude model, carrier transport is described by a single parameter, namely the carrier mobility μ. At low-electric fields, and in the absence of external magnetic fields, μ is the constant of proportionality between the average carrier velocity $\mathbf{v}$ and an external electrical field $\mathbf{E}$:

$$\mathbf{v} = \mu \cdot \mathbf{E} \qquad 15.1$$

μ is indirectly proportional to the carrier's effective mass m^* and directly proportional to the averaged transport scattering time $<\tau_t>$:

$$\mu = \frac{e}{m^*} \langle \tau_t \rangle \qquad 15.2$$

where e is the electron charge. Within the limits of the wave-vector-independent relaxation time approximation, $1/<\tau_t>$ is the sum of all reciprocal scattering times associated with all possible scattering mechanisms (Mathiessen's rule). Evidently, the mobility is dominated by the mechanism with the shortest scattering time.

The main scattering mechanisms in elemental bulk semiconductors that contribute to $<\tau_t>$ are scattering at acoustic and optical phonons, as well as scattering at ionized and neutral impurities. In $Si_{1-x}Ge_x$ layers random alloy scattering contributes as another independent mechanism. In low-dimensional carrier systems, the presence of heterointerfaces (including the Si/SiO_2 and the Si/high-k-dielectric systems) leads to interface scattering,[13] which itself comprises scattering at interface roughness, interface charges, and interface-related phonons. Strain will have influence on most of these scattering mechanisms, because the strain-induced changes in the valence and conduction band structure affect the relative importance of intra- and intervalley scattering events.[14]

Another mechanism that is indirectly related to strain is scattering at dislocations that penetrate the active layers.[15] Such dislocations are often unavoidable byproducts of strain engineering via metamorphic layers, i.e., virtual substrate layers that are on purpose designed to become strain relaxed via misfit dislocations.[10] In addition, changes of the band structure are usually concomitant with modifications of the effective mass m^*, which directly enters the mobility term in eq. 15.1. Hence, the modeling of transport in a strained-layer system is a formidable task, which has been addressed by several groups, [15–18] as will be described in some detail in the following sections.

Ample experimental and theoretical data exist for bulk Si, and, to some extent, for bulk Ge. The available models work best for unstrained n-type bulk material at temperatures above 100 K and for rather small doping concentrations below some 10^{17} cm^{-3}. Under these conditions intra- and intervalley lattice scattering is dominating. With phonon freeze-out at cryogenic temperatures, ionized impurity scattering becomes the limiting mechanism, whereas the influence of neutral impurity scattering remains moderate.

For higher doping concentrations and/or lower temperatures, phenomenological expressions have been derived, which are based on a combination of physical models and fits to experimental data. This is especially true for holes, which, due to the complexity of the valence band structure, are hard to describe by physical models for larger variations of doping concentration or temperature. A concise phenomenological model description was published by Klaassen, who treated both electrons and holes in unstrained bulk Si over very wide ranges of doping concentrations and temperatures.[18]

As a reference for the subsequent sections, the experimental room temperature electron and hole mobilities, as well as the effective masses, are listed in Table 15.1 for undoped, unstrained bulk Si, Ge, and, when available, SiGe. The intrinsic mobilities are of special interest, because they represent the limiting case of lattice scattering alone, as neither neutral nor ionized impurities are present. Hence, these are the maximum values achievable at room temperature as long as the effective mass of the system is not affected.

The values in Table 15.1 reveal that both electron and hole mobilities are significantly larger in Ge, which is mainly a consequence of the smaller

Table 15.1 Essential band and transport parameters of unstrained Si, Ge, and SiGe (From Ref. 12)

	Si	Ge	$Si_{1-x}Ge_x$
Effective electron mass (m_0)	$m_l^* = 0.92$	$m_l^* = 1.59$	$m_l^* \approx 0.92$ ($x \le 0.85$) $m_l^* \approx 1.59$ ($x > 0.85$)
	$m_t^* = 0.19$	$m_t^* = 0.08$	$m_l^* \approx 0.19$ ($x \le 0.85$) $m_l^* \approx 0.08$ ($x > 0.85$)
Effective hole mass (m_0)	$m_{hh}^* = 0.54$ $m_{lh}^* = 0.15$ $m_{so}^* = 0.23$	$m_{hh}^* = 0.33$ $m_{lh}^* = 0.043$ $m_{so}^* = 0.095$	
Electron mobility (cm^2 V^{-1} s^{-1})	1450	3900	$1450 - 4325x$ ($0 \le x < 0.3$)
Hole mobility (cm^2 V^{-1} s^{-1})	450	1900	$450 - 865x$ ($0 \le x < 0.3$)
Energy gap at 300 K (eV)	1.12 (Δ)	0.66 (L)	$1.12 - 0.41x + 0.008x^2$ ($x \le 0.85$) $1.86 - 1.2x$ (L-like for $x > 0.85$)
Spin-orbit splitting Δ_0 (eV)	0.044	0.29	$0.44 + 0.246x$

m_0 = free electron mass

effective masses. Especially the hole mobility of Ge is worth mentioning, since it is higher than in any other group-IV or III–V semiconductor, and, moreover, almost matches the intrinsic electron mobility of Si.

While experimental data for the carrier mobilities in Si and Ge are available, and satisfactory models exist, the experimental and theoretical situation concerning the $Si_{1-x}Ge_x$ alloys is more rudimentary. It is clear from the above discussion that not only alloy scattering has to be considered as an additional mechanism, but also the other mechanisms are modified by the variations of the band structure and of the phonon spectra. The most prominent and experimentally confirmed example is the transition from a Si-like to the Ge-like conduction band, which occurs in unstrained alloys at $x = 0.85$ (Table 15.1). This transition leads to a kink in the electron mobility as a function of the composition that essentially follows the change of the effective mass.[19]

Besides the electron mobilities, the hole mobilities in $Si_{1-x}Ge_x$ alloys are of interest. In unstrained bulk alloys early experiments found a U-shaped mobility behavior, i.e., the lowest mobilities were found at intermediate composition.[20] This trend was confirmed by calculations,[16, 21] which predict a strong influence of strain in pseudomorphic $Si_{1-x}Ge_x$ layers on Si substrates. A more detailed discussion of the experimental results, and the controversial theoretical predictions for strained $Si_{1-x}Ge_x$ layers, will be given in Section 15.4.2 below, which deals with the properties of 2D hole gases in modulation-doped structures.

15.3 Strain engineering

As has been shown in the preceding section, the strain-dependence of the band structure in Si/SiGe heterostructures plays a crucial role for the transport properties of electrons and holes. Generally, an arbitrary strain tensor can be separated into a hydrostatic and a uniaxial component:

$$\vec{\varepsilon} = \begin{pmatrix} \varepsilon_{\parallel} & & \\ & \varepsilon_{\parallel} & \\ & & \varepsilon_{\perp} \end{pmatrix} = \varepsilon_{\parallel} \cdot \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} + \begin{pmatrix} 0 & & \\ & 0 & \\ & & \varepsilon_{\perp} - \varepsilon_{\parallel} \end{pmatrix} \qquad 15.3$$

where $\varepsilon_{\parallel} = \left(\frac{a_{\parallel}}{a} - 1\right)$ and $\varepsilon_{\perp} = \left(\frac{a_{\perp}}{a} - 1\right)$ are the strain components in the growth plane and growth direction, respectively, and the a-terms are the corresponding components of the lattice constant. The hydrostatic strain component affects the band energies, but preserves the crystal symmetry, and thus the band degeneracies. The uniaxial component, on the other hand, reduces the symmetry of the crystal, and thus can lift the band degeneracies. The weighted average band energy remains invariant under uniaxial strain. The amount of the energetic shifts and splittings of the respective band depends on the corresponding deformation potentials and the amount and direction of strain for a given band and its position within the Brillouin zone.[23]

For a complete description of the band alignment problem the relative positions of the bands in the two heteromaterials without strain are required. These depend on the compositions of the two heteromaterials in contact, and on the orientation of the interface. Several groups have addressed this problem with pseudopotential approaches[16,22] and density functional theory.[23,24] The latest results employ *ab initio* calculations with a dynamically screened interaction correction[25] (GW approximation[26]), which are known for their accuracy regarding band gap predictions. Based on the *ab initio* results, fitting parameters were then extracted for a much faster, but still highly accurate 30-level $\boldsymbol{k}\cdot\boldsymbol{p}$ model.[25]

To qualitatively show the basic effects of strain on the band structure, we use here an analytical expression for the strain dependence of the weighted average valence band offset $\Delta E_{\mathrm{v}}^{\mathrm{av}}$ at $\Gamma_{25'}$, which was shown to depend only weakly, and in a linear fashion, on strain, and to be widely independent of the crystal orientation.[23] This means that $\Delta E_{\mathrm{v}}^{\mathrm{av}}$ can be exploited as a reference energy for the intrinsic band alignment in the Si/SiGe heterosystem. From the calculated values of $\Delta E_{\mathrm{v}}^{\mathrm{av}}$ in References 23 and 24, Rieger and Vogel extracted a relation that expresses $\Delta E_{\mathrm{v}}^{\mathrm{av}}$ as a function of the composition x of a strained film on a metamorphic (i.e. unstrained) substrate of composition x_{s}:

$$\Delta E_{\mathrm{v}}^{\mathrm{av}}(\mathrm{eV}) = (0.47 - 0.06x)(x - x_{\mathrm{s}}) \qquad 15.4$$

Calculating the band shifts and splittings in the strained case is then a straightforward task. Here we give the results for the most relevant situation of (001) oriented substrates and biaxial in-plane strain. Other strain situations can easily be constructed by making use of the additivity and transitivity of the band alignment.[23] Also, numerical results for band variations and splittings on (001), (011) and (111) oriented substrates can be found in Ref. 25.

The hydrostatic strain component leads to a change in the band gap according to:

$$\Delta E_{\mathrm{g}} = \left(\Xi_{\mathrm{d}} + \frac{1}{3}\Xi_{\mathrm{u}} - \alpha\right)(\varepsilon_{\perp} + 2\varepsilon_{\parallel}) \qquad 15.5$$

where $\Xi_{\mathrm{d}} + \frac{1}{3}\Xi_{\mathrm{u}} - \alpha$ is the deformation potential difference for the fundamental band gap in a Si-like band structure.[12]

The sixfold degenerate Δ_6 conduction band splits into the four in-plane valleys (Δ_4) and the two valleys along the growth direction (Δ_2) according to:

$$\Delta E_c(\Delta_2) = +\frac{2}{3}\,\Xi_u\ (\varepsilon_\perp - \varepsilon_\parallel) \qquad 15.6$$

$$\Delta E_c(\Delta_4) = -\frac{1}{3}\,\Xi_u\ (\varepsilon_\perp - \varepsilon_\parallel) \qquad 15.7$$

The valence band at the Γ point splits into heavy holes (hh), light holes (lh) and the spin-orbit split hole band (so):

$$\Delta E_v(\mathrm{hh}) = +\frac{1}{3}\,\Delta_0 - \frac{1}{2}\,\delta E_{001} \qquad 15.8$$

$$\Delta E_v(\mathrm{lh}) = -\frac{1}{6}\,\Delta_0 + \frac{1}{4}\,\delta E_{001} + \frac{1}{2}\left(\Delta_0^2 + \Delta_0\delta E_{001} + \frac{9}{4}\,\delta E_{001}^2\right)^{1/2} \qquad 15.9$$

$$\Delta E_v(\mathrm{so}) = -\frac{1}{6}\,\Delta_0 + \frac{1}{4}\,\delta E_{001} - \frac{1}{2}\left(\Delta_0^2 + \Delta_0\delta E_{001} + \frac{9}{4}\,\delta E_{001}^2\right)^{1/2} \qquad 15.10$$

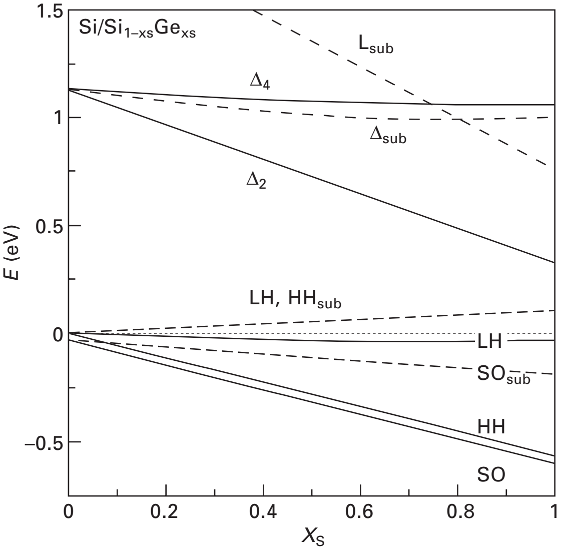
where $\delta E_{001} = 2b(\varepsilon_\perp - \varepsilon_\parallel)$, with the valence band deformation potential b, and Δ_0 is the spin-orbit splitting parameter (Table 15.1).

Experimentally, only deformation potential differences between two bands are accessible. Therefore, calculated deformation potentials[23–25] and linear combinations for different compositions are usually employed.[10,12]

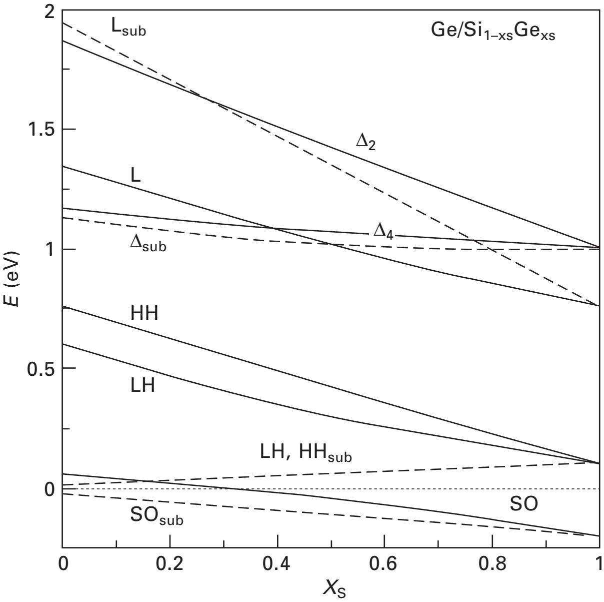
Since Si/SiGe heterostructures have a type-II band alignment under almost all conditions,[9,27] for electron confinement only the originally sixfold degenerate Si-like Δ_6 conduction band, and for hole confinement only the originally twofold degenerate heavy hole and light hole valence band states in SiGe are relevant. The only important exception is the use of SiGe stressors in Si MOSFETs, which are positioned outside the active Si region, and thus affect the Si valence band via uniaxial strain splitting and effective mass modification. This situation is treated in Section 15.5.

For the following sections we concentrate on biaxial strain in the (001) plane, which is introduced by a metamorphic $\mathrm{Si}_{1-xs}\mathrm{Ge}_{xs}$ substrate.[10] The variations of the relevant bands as a function of x_s are plotted in Figs 15.1 and 15.2 for a pure Si and a pure Ge channel, respectively.[9] Besides the aforementioned Δ conduction bands and the hh and lh bands, also the L band, which is the conduction band minimum in unstrained Ge, and the spin split-off valence band (so) are plotted in the two figures. The solid lines refer to the respective bands in the strained Si or Ge layer; the dashed lines refer to the bands in the unstrained substrate with composition x_s.

The most important influence of strain is the splitting of the valence and conduction bands, as described in eqs 15.6–15.10. Under tensile in-plane strain, as in Fig. 15.1, Δ_2 is lowered and defines the global conduction band minimum, whereas Δ_4 is raised above the Δ minimum of the $\mathrm{Si}_{1-xs}\mathrm{Ge}_{xs}$ substrate.



15.1 Variation of the relevant conduction and valence bands of a strained Si layer on an unstrained $Si_{1-xs}Ge_{xs}$ substrate as a function of the substrate composition x_s. (From Ref. 9)



15.2 Variation of the relevant conduction and valence bands of a strained Ge layer on an unstrained $Si_{1-xs}Ge_{xs}$ substrate as a function of substrate composition x_s. (From Ref. 9)

The in-plane transport mass of the Δ_2 valley is the transversal effective mass $m_t = 0.19m_0$, which is the lowest possible electron mass. Both theoretical and experimental treatments show that m_t is hardly affected by strain.[10,16] For the case of a compressively strained Ge channel on a $Si_{1-xs}Ge_{xs}$ substrate (Fig. 15.2) both hole bands are raised with respect to the substrates, and, in addition, they split in favor of the hh band. The hole transport mass is strongly strain dependent and also depends on the carrier concentration in the channel due to band filling of the severely non-parabolic band.[10,16] These effects are discussed from an experimental point of view on page 423.

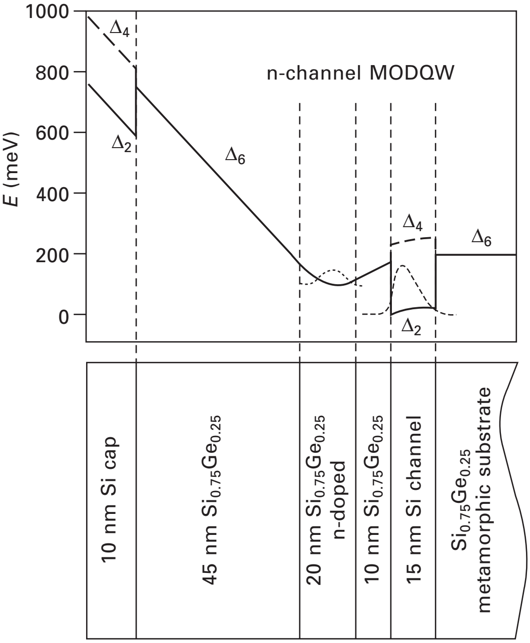
15.4 Low-dimensional transport

15.4.1 Two-dimensional electron gases (2DEG)

Low-temperature mobility and metal-to-insulator transition

The strain-dependent band alignment in the Si/SiGe heterosystem leads to a finite conduction band offset only in the case of a Si-rich channel under tensile in-plane strain.[23] Such conditions were first implemented by Jorke and Herzog,[28] who employed a fully strain-relaxed $Si_{1-x}Ge_x$ film on Si(001) as a metamorphic substrate for the Si channel. Electrons were supplied from an n-doped SiGe layer above (in the growth direction) the channel that was separated from the channel by an undoped SiGe spacer layer. This basic modulation-doping (MOD) layer sequence is still employed for high-mobility n-channel structures, such as those discussed in the following subsections. Recent refinements have been concentrated mainly on improved defect control in the metamorphic substrate and on optimized temperature and composition profiles during growth. Also, capping layers with reduced doping concentration, usually terminated by a thin Si layer to protect the SiGe layers against oxidation, were introduced for control of the carrier concentration via well-defined Schottky gates.

A typical layer sequence and the self-consistently derived potential variation across it are depicted in Fig. 15.3 for an n-type modulation-doped Si/SiGe quantum well structure (MODQW) with pure Si channel. Proper design of a MODQW structure requires adjustments of the layer thicknesses, the compositions and strain states, and the doping concentrations. The most basic design rule requires that under operational conditions all dopant atoms should be ionized, whereas all free carriers should be confined to the channel region. This condition defines the integral doping concentration for a given layer sequence and a given surface potential, but still leaves ample design freedom for optimizing an MODQW structure for different purposes. High mobilities, for example, require thick spacer layers to reduce Coulomb scattering at the ionized impurities of the doping supply layer. Because of the linear potential drop across the undoped spacer, charge transfer from


1000
800
600
400
200
0
E (meV)
Δ4
Δ2
Δ6
n-channel MODQW
Δ4
Δ6
Δ2
10 nm Si cap
45 nm Si0.75Ge0.25
20 nm Si0.75Ge0.25 n-doped
10 nm Si0.75Ge0.25
15 nm Si channel
Si0.75Ge0.25 metamorphic substrate

15.3 Conduction band variation over an n-type Si/SiGe MODQW structure on a metamorphic $Si_{0.75}Ge_{0.25}$ substrate. The Δ_6 conduction band splits only in the strained Si layers into a lower Δ_2 and a higher Δ_4 band. The wave functions of the lowest states in the Si channel and in the doping supply layer are indicated as dashed lines.

the supply layer into the channel becomes less and less efficient as the spacer is increased. In contrast, device applications require a high channel conductance, for which the product of the density of carriers in the channel and their mobility has to be optimized.

The reported low-temperature electron mobilities in Si/SiGe MODQWs have reached peak values of 660,000 cm^2 V^{-1} s^{-1},[29,30] with mobilities between 200,000 and 500,000 cm^2 V^{-1} s^{-1} having been routinely achieved by several laboratories over the last 10 years.[31–34] While the carrier mobility is an important figure-of-merit at the higher carrier concentrations typically found in transistors, it is less relevant at low carrier concentrations, where local potential fluctuations begin to dominate transport. This was shown by Wilamowski *et al.*, who, in a conduction electron spin resonance (CESR) experiment, measured the density of states at the Fermi level, $D(E_F)$, as a function of the carrier concentration n.[35] Rather than finding a constant $D(E_F)$, as would be expected for an undisturbed 2DEG, they found a value that decreased gradually with decreasing n. Since for weakly interacting electrons the Thomas–Fermi screening wave vector is proportional to $D(E_F)$, screening becomes evidently less efficient with decreasing *n*. This behavior was interpreted by potential fluctuations δV, which lead to tail states that make screening less efficient, and, in a positive feedback, cause an increase of the potential fluctuations. Since both δV and E_F are experimentally accessible by CESR experiments, it was indeed found that at low enough *n,* the fluctuation energy $e\delta V$ first becomes larger than E_F and finally diverges. This behavior causes carrier localization at the critical value of $n_s = n_{MIT}$, where MIT stands for the metal-to-insulator transition. In contrast to MOSFETs, the transport properties of which are determined by large angle scattering at interface charges, carriers in MODQW structures are mainly scattered by small angles at the long-range Coulomb potential of remote dopants. For the latter condition, comparable length scales for potential fluctuations and carrier localization are expected, with a lower, experimentally determined[35] limit of about 1 μm.

In this picture, puddles of mobile electrons are confined to shallow long-range potential minima, and the landscape of the potential fluctuations determines the transport properties at low carrier concentrations. It is evident that the carrier mobility is not a very meaningful quantity at such low carrier concentrations, where mobile and localized carriers undergo phase separation, and transport is mainly determined by percolation between puddles of mobile carriers. The potential fluctuations themselves and their characteristic length scale were mainly attributed to dopant fluctuations in the doping supply layer,[35] and to a lesser extent, to interface scattering, and, if not well controlled, to background doping. It is therefore not surprising that samples grown under different conditions can vary quite significantly in their n_{MIT} values, even if they show high mobilities at higher *n*. For instance, the sample with the highest reported mobility[29] had an n_{MIT} value of $\approx 10^{11}$ cm^{-2}, whereas the Si/SiGe sample with the as yet lowest reported n_{MIT} value of 3.2×10^{10} cm^{-2} had a mobility of just around 200,000 cm^2 V^{-1} s^{-1}.[36] The sample used in Ref. 35 for the screening investigations was somewhere in between, with $\mu = 200{,}000$ cm^2 V^{-1} s^{-1} and $n_{MIT} = 7\cdot 10^{10}$ cm^{-2}.

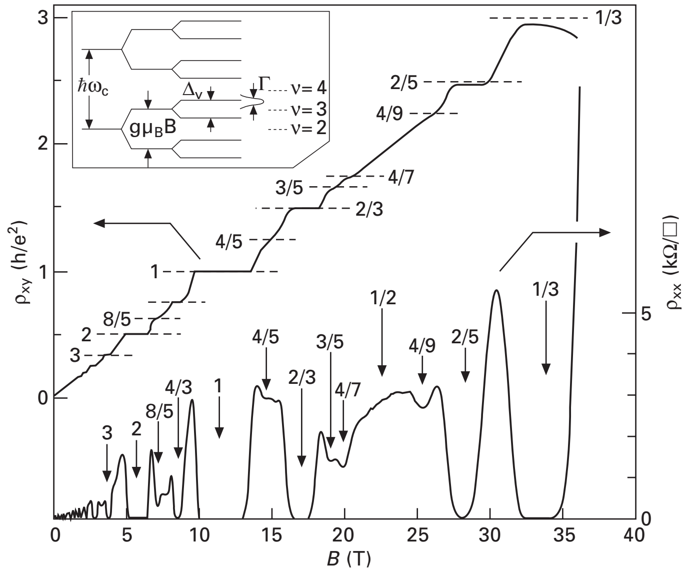
We recently tried[37] to achieve higher mobilities by increasing the Ge content in the barrier to $x = 0.3$, and thus increase both the conduction band offset and *n*, as compared to our older samples with $0.2 < x < 0.25$, which were employed for the studies in Refs 35 and 36. Indeed, the mobilities went up by a factor of about 2, to $\mu = 420{,}000$ cm^2 V^{-1}s^{-1} at $n_s = 4.5 \times 10^{11}$ cm^{-2}, but simultaneously, also n_{MIT} went up to a disappointing $n_{MIT} = 2.7 \times 10^{11}$ cm^{-2}.[38] Evidently, the potential landscape that leads to a rapid loss of the apparent carrier mobility with decreasing *n* is very efficiently screened at higher *n*. This preliminary result might be indicative of the role of variations in the potential barrier caused by local composition fluctuations. Indeed, for random alloys, composition fluctuations should be proportional to $[x(1-x)]^{-1}$, which becomes maximum at a composition of $x = 0.5$.[39]

Clearly, additional experiments are required to find growth conditions that simultaneously optimize both μ and n_{MIT}. It should, however, be mentioned that both types of samples, i.e. those that were optimized for high μ and those optimized for low n_{MIT}, were successfully employed both for experiments regarding the MIT itself,[29,35,36,40] and for transport experiments in the (fractional) quantum Hall regime, which will be summarized in the next section.

Quantum Hall effect

The integer quantum Hall effect (IQHE) was originally discovered on 2DEGs in Si MOSFETs,[41] but subsequent research was mainly concentrated on III–V heterostructures with their much superior mobilities. In particular, the discovery[42,43] of the fractional quantum Hall effect (FQHE) would not have been possible on the basis of MOSFETs with their mobility limiting, large-angle interface scattering properties. Meanwhile, the availability of high-mobility Si/SiGe heterostructures has strongly reduced the performance gap to the III–V semiconductors. In addition, electrons in strained Si channels differ from their III–V counterparts because of the twofold degeneracy of the Δ_2 valleys in the growth direction. The relevance of the valley degeneracy has been a major concern regarding the spin coherence of 2DEGs in strained Si channels,[44,45] and it was also not clear to what extent it would affect the many-body description of the FQHE.

To clarify these basic problems, the QHE was studied in Si/SiGe heterostructures by several groups, who reported indications of FQHE states measured on a variety of samples from different laboratories.[46–50] The most concise experiments so far were performed in the group of D. C. Tsui, who employed magnetic fields B of up to 45 T and temperatures down to 30 mK.[51] The investigated sample had a mobility of 250,000 cm^2 V^{-1} s^{-1} and an $n_{\mathrm{MIT}} < 5 \times 10^{10}$ cm^{-2}. Figure 15.4 shows an overview of longitudinal and lateral resistivities, ρ_{xx} and ρ_{xy}, respectively, in the range $0 < B < 40$ T at 30 mK. The Shubnikov-de-Haas oscillations are resolved down to a filling factor of $\upsilon = 36$. The first odd IQHE state appears at $B = 1$ T and $\upsilon = 11$. At this magnetic field, the splitting Δ_{v} between the Δ_2 valleys was estimated to be about 26 μeV (corresponding to a thermal energy of 0.3 K). For $\upsilon < 1/3$ the sample enters an insulating state. Around $\upsilon = 1/2$ the principal FQHE states are observed at $\upsilon = \frac{2}{3}, \frac{3}{5}$ and $\frac{4}{7}$; and the two-flux series is observed at $\upsilon = \frac{4}{9}, \frac{2}{5}$ and $\frac{1}{3}$. These results demonstrate that the basic concept of the composite fermion (CF) model[52] remains valid, despite the twofold valley degeneracy. Careful mapping of the energy gaps of the observed FQHE states revealed quite surprisingly that the CF states assume their own valley degeneracy, which appears to open a gap proportional to the effective magnetic field B^* of the respective CF state, rather than being proportional to the absolute B



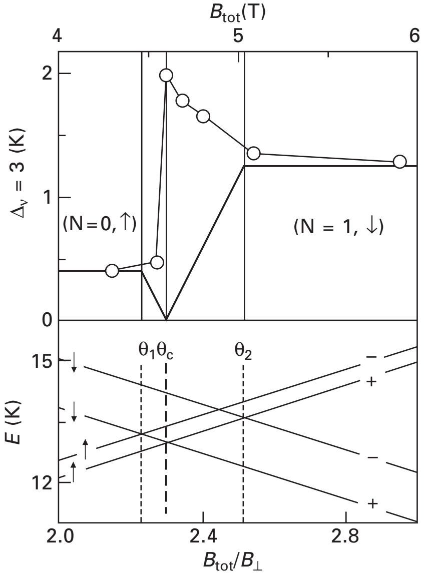
15.4 Diagonal resistivity ρ_{xx} and Hall resistivity ρ_{xy} of the 2DEG in a strained Si quantum well at $T = 30$ mK. Major fractional quantum Hall states are marked by arrows. The inset shows the Landau level diagram. Filling factors are labeled υ; the level broadening is denoted by Γ. (From Ref. 50)

field.[53] For the CF states the valley degeneracy therefore plays a different role than the spin degeneracy, the opening gap of which is proportional to B, and thus does not play a role at the high magnetic fields at which FQHE states are typically observed.

To gain further insight into the properties of the degenerate Δ_2 valleys, several groups performed coincidence experiments in tilted magnetic fields. These experiments make use of the fact that the Landau levels are separated by the cyclotron gap, $E_C = \hbar e B_\perp / m^*$, which depends only on the magnetic field component $B_\perp$ perpendicular to the 2DEG. On the other hand, Zeeman spin splitting, $E_Z = g^* \mu_B B$, is proportional to the total magnetic field B, i.e. independent of the orientation of B with respect to the 2DEG. Here g^* and μ_B are the effective g-factor and the Bohr magneton, respectively. Hence, when tilting the magnetic field out of the direction normal to the 2DEG, the spin splitting becomes enhanced relative to the Landau splitting, and coincidences occur at well-defined tilting angles, where spin and Landau levels cross. Since the valley degeneracy is also lifted in magnetic fields, the behavior of the valleys can be sensitively studied in the coincidence regime of odd IQHE states, for which the Fermi level lies between two valley states.[54]

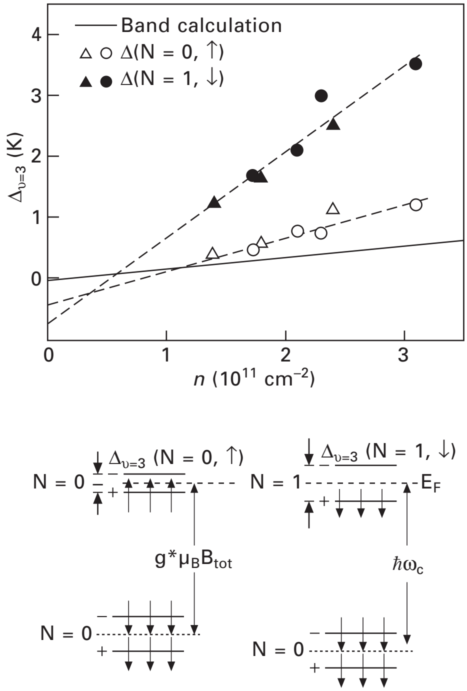
Lai and coworkers performed such coincidence experiments at odd integer filling factors of $\upsilon = 3$ and $\upsilon = 5$,[55] and, for comparison at the even integer filling factors $\upsilon = 4$ and 6.[56] In agreement with earlier experiments, they observed that outside the coincidence regime of odd integer filling factors the valley splitting does not depend on the in-plane component of the magnetic field. However, the valley splitting is significantly different (by up to a factor of 3 for $\upsilon = 3$) in the regions right and left of the coincidence regime. Moreover, the valley splitting shows a pronounced anomaly inside the coincidence regime, where it becomes enhanced rather than suppressed, as would have been expected in a single particle picture (Fig. 15.5). This anomaly was shown to be missing in the coincidence regime of even filling factors.

To elucidate the origin of this unexpected behavior, the dependence of the valley splitting on the carrier density n was investigated in the range below ($\Delta_3(N = 0, \uparrow)$ state) and above ($\Delta_3(N = 1, \downarrow)$ state) the $\upsilon = 3$ coincidence in Ref. 56. Here N is the Landau level index, and ($\downarrow, \uparrow$) are the two spin orientations. Note that we use here the common nomenclature of the $\downarrow$ spin state being anti-parallel to $\boldsymbol{B}$, and therefore defining the energetically lower Zeeman state in the Si/SiGe material system with its positive g^*; in Refs



15.5 Upper panel: measured $\Delta_{\upsilon=3}$ gap (circles) close to the $\upsilon = 3$ coincidence region. The single particle gap calculated from a Landau fan diagram is shown as a solid line. Lower panel: Landau fan diagram in tilted B fields, with $B_{tot}/B_\perp$ on the x-axis. The energy levels are labeled with the Landau level index N, the spin orientation ($\downarrow, \uparrow$) and the valley index (+, –). The three crossing levels are labeled θ_1, θ_2 and θ_C. The latter is the usual coincidence angle, where level crossing occurs at the Fermi level. (From Ref. 54)

55 and 56, spin labeling was reversed. A linear n-dependence was found for either configuration, though with significantly different slopes (Fig. 15.6). Moreover, both slopes are higher than that of the bare valley splitting predicted by a band calculation at $B = 0$.[56] The configurations below and above the $\upsilon = 3$ coincidence differ in both the Landau level indices and the spin orientation. Therefore, the origin of the different n-dependencies could simply represent the different exchange-correlation energies of the $N = 0$ and $N = 1$ Landau levels. This approach, however, turned out to be inconsistent with the experimental n-dependence. A consistent interpretation is based on electron–electron interaction, the energy contribution of which is comparable to the Landau and spin-splitting energies in the coincidence regime. For electron–electron interaction the spin state of the highest occupied level is relevant, taking into account that the lower two levels are both ($N = 0$, ↓)



15.6 Upper frame: density dependence of the valley splitting at $\upsilon = 3$. Empty symbols stand for $\Delta_3(N = 0, \uparrow)$, filled symbols for $\Delta_3(N = 1, \downarrow)$. Dashed lines are linear fits to the data that extrapolate to finite values at zero density. The solid line shows the calculated single-particle valley splitting. Lower frame: schematic arrangement of the relevant energy levels near the Fermi level E_F, including the two lowest ($N = 0$, ↓, + −) states. From the spin orientation in the three occupied levels it becomes clear that the Pauli exclusion principle diminishes screening of the ($N = 1$, ↓) states. (From Ref. 55) states that differ only in their valley quantum number (labeled + and - in Figs 15.5 and 15.6). Thus, below the coincidence regime, the electrons of the two lower states have opposite spin with respect to the highest occupied (N = 0, ↑) state (Fig. 15.6). Screening of the Coulomb interaction is therefore efficient, and the *n*-dependence is closer to the bare valley splitting. Above the coincidence regime, however, screening by the two lower states becomes diminished by the Pauli exclusion principle, because now all three states are spin-down states. Consequently, the $\Delta_3(N = 1, \downarrow)$ gap is greatly enhanced over the bare valley splitting (Fig. 15.6).

Coincidence experiments have also been used to study quantum Hall ferromagnetism (QHF) in strained Si channels with Δ_2 valley degeneracy. QHF can be expected when two energy levels with different quantum indices become aligned and competing ground state configurations are formed. If in such a case the magnetic order of the system becomes anisotropic with an easy axis, then the system behaves similar to an Ising ferromagnet.[57] In particular, in the strong electron–electron interaction regime QHF may occur, when two levels with opposite spin (or quasi-spin) states cross each other.

Strong indications for QHF in a strained Si/SiGe heterostructure were observed[58] around $\upsilon = 3$ under the same experimental coincidence conditions as the aforementioned experiments regarding anomalous valley splitting. Under these conditions a hysteretic magnetoresistance peak was observed, which moves from the low field to the high field edge of the QHE minimum as the tilting angle of the magnetic field passes through the coincidence angle.

More detailed studies were reported by the group of T. Okamoto, who employed a sample with a mobility of 480,000 $\text{cm}^2\ \text{V}^{-1}\ \text{s}^{-1}$.[59] They measured the resistance along a Hall bar in a magnetic field that was tilted away from the normal to the 2DEG by an angle Φ. The in-plane field component was rotated with respect to the current direction of the Hall bar by an azimuth angle φ, with $\varphi = 0°$ standing for the in-plane magnetic field component being along the current direction. With Φ adjusted to the coincidence angle Φ_c, the longitudinal resistivity ρ_{xx} was measured as a function of φ. Again coincidence of the $(N = 0; \uparrow)$ and the $(N = 1; \downarrow)$ levels was investigated. The authors found a resistance peak at Φ_c, which was especially high around $\upsilon = 4$. Moreover, they found a large in-plane anisotropy, with the peak height for $\varphi = 0°$ being much higher than for $\varphi = 90°$. A very similar behavior had been observed before by Zeitler *et al.* at higher magnetic fields on samples with somewhat lower mobilities.[60] Zeitler *et al.* interpreted their results in terms of a unidirectional stripe phase developing at low temperatures in a direction perpendicular to the in-plane magnetic field component. Such a stripe phase was also assumed by Okamoto *et al.*, who assigned the stripes to the domain structure of Ising ferromagnets. The latter postulation is based on the pronounced hysteresis of the resistance anomaly at temperatures between

50 and 300 mK. Above 300 mK the resistance peak vanishes rapidly, which is indicative of the collapse of the Ising ferromagnetic domain structure. Although it is not entirely clear what role the twofold valley degeneracy in the strained Si channels plays for the QHF, Okamoto *et al.* conclude from the measured temperature dependence that it cannot dominate the breakdown of Ising ferromagnetism.

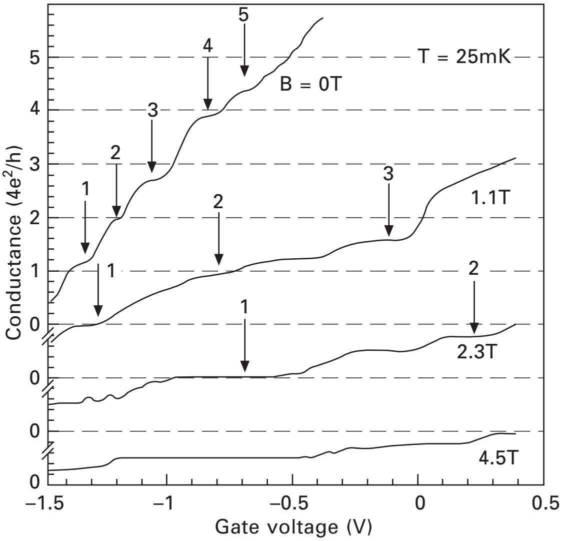
Quantum point contacts

The low-temperature results in the preceding section on 2DEGs in strained Si channels were performed on relatively large (several 100 μm long) structures designed for macroscopic Hall measurements. They result in valley splittings that are, even if enhanced in the coincidence regime at an odd integer filling factor, on the order of a few 10 μeV (corresponding to a thermal energy k_BT of just a few K) in magnetic fields around 1 T. It was always clear that the valley splitting can be lifted in high enough magnetic fields, because of its experimentally confirmed linear increase with B;[53] however, the experimental results gathered on macroscopic samples indicated that the valley splitting is smaller than the spin splitting. Recently, Goswami *et al.*[61] and, in an extensive follow-up work, McGuire *et al.*[62] provided experimental evidence for strongly enhanced valley splitting in a rectangular quantum well, under conditions of strong lateral carrier confinement. Generally, the valley splitting should be lifted in the presence of abrupt potential walls, provided they induce a strong enough coupling of the two valleys.[63] Quantitatively, the influence of the quantum well thickness on the valley splitting was studied by a tight-binding simulation,[64] which predicted (i) valley splitting at zero electric and magnetic field, (ii) oscillations of the zero field splitting with well thickness, and (iii) an overall decay of the splitting with the third power of the QW thickness. Thus, in order to obtain well-defined valley splitting, the QW has to be sufficiently narrow, and its thickness has to be maintained over the extension of the electron wave function close to atomic layer precision. Of course, atomic layer control of the QW thickness is virtually impossible in large area samples, but in a laterally confined system containing only a few electrons, chances are high that the electron wave function, which then probes only a very small area of the confining heterointerfaces, encounters a sufficiently low amount of QW thickness fluctuations. Consequently, quantum point contacts (QPC) were employed in Refs 61 and 62, which typically provide lateral confinement by the adjustable electrostatic fields of Schottky gates.

Többen *et al.* were the first to fabricate QPCs in the 2DEG of a Si/SiGe MODQW structure by employing electron beam lithography to define a Pd split-gate on a Hall bar.[65] The 2DEG in their sample had a 2D carrier density of $n = 4.5 \times 10^{11}$ cm^{-2}, and a mean free path of $l_{mfp} = 1.3$ μm at 25 mK.

The authors observed as a function of the split-gate voltage conductance plateaus in units of $4e^2/h$ (Fig. 15.7), which result from 1D quantization in an electron system with spins and valleys being both twofold degenerate, i.e. $g_s = g_v = 2$. At rather modest magnetic fields applied perpendicular to the 2DEG, both spin and valley degeneracy became lifted.[66] Already at 2.3 T, first indications for conductance quantization in units of e^2/h were observed, which became fully resolved at 4.5 T.[65] Fully lifted degeneracies were also reported more recently by Scappucci *et al.* in 1D transport realized by a wrap-around gate configuration on a narrow Si/SiGe mesa structure.[67]

Now, Goswami *et al.*[61] and McGuire *et al.*[62] performed very careful new experiments on split-gate QPCs, with the aim of quantitatively determining the valley splitting as a function of the confining electrostatic field, and of a magnetic field perpendicular to the 2DEG. In their model descriptions employed to extract the valley splitting from the conductance behavior of the QPC, the QW barriers were assumed to contain evenly spaced monoatomic height steps to model their samples, which were grown on substrates miscut by 2° from the ideal (001) orientation. Similar, albeit not periodic, QW thickness fluctuations can also be expected in most other high-mobility samples grown so far, because of the dislocation-induced cross-hatch pattern typically encountered in relaxed SiGe buffer layers with graded composition profiles.[9,10]



15.7 Conductance in units of $4e^2/h$ of a quantum point contact (QPC) with and without a perpendicular magnetic field. The lifting of the spin and valley degeneracies is clearly resolved at B = 4.5 T. (From Ref. 64)

The new experiments revealed that the valley splitting can indeed be controlled by lateral carrier confinement. It was shown that both an electrostatic confining potential, as in the case of a QPC, and a perpendicular magnetic field, which leads to lateral confinement on the order of a cyclotron orbit, can provide such confinement. However, quantitative extraction of the valley splitting turned out to be rather complex. A total of three different model descriptions and extensive fitting routines with elaborate error analyses were employed in Refs 61 and 62 to deduce absolute values for the valley splitting. Despite relatively large variations between the different model descriptions, the fitting routines led for the lowest QPC energy level to a zero-magnetic field valley splitting on the order of 0.5 meV (corresponding to 6 K), which increases to about 1.5 meV (17 K) at $B = 8$ T. The first excited QPC state appears to behave in a much more complex manner, showing a non-monotonic valley splitting ranging from 0 to about 1 meV in the same magnetic field range. At such high magnetic fields, where magnetic confinement becomes effective, comparable valley splittings in the meV range were earlier reported by Khrapai *et al.*[53] and Wilde *et al.*[68] The latter group performed deHaas–van Alphen experiments that allow for a direct extraction of the valley splitting at high magnetic fields.

Evidently, and despite great efforts in the recent past, valley splitting will remain a controversial topic for the time being. Nevertheless, the new experiments on QPC may have identified an important contribution to the apparent valley splitting, namely the effect of QW roughness on an atomic scale, which diminishes the barrier-induced coupling, and thus the splitting, of the two originally degenerate Δ_2 valleys. There is now some confidence that in laterally confined few-electron systems valley splitting is much more pronounced, and valley splitting on the order of meV can be expected for the QPC ground state even in the absence of an external magnetic field. These results are indeed promising for applications based on spin manipulation, because such systems will require confined few-electron states with long spin coherence times that are not jeopardized by intervalley scattering.

Room-temperature electron mobility

While high mobilities at low-temperatures are a precondition for basic studies of the 2DEG properties in the Si/SiGe heterosystem, the room-temperature (RT) performance is an important figure-of-merit for device applications. RT mobilities significantly higher than the bulk mobility in unstrained and undoped Si are to be expected in tensile strained Si channels, because only the Δ_2 valleys with the lowest possible transport mass in Si, $m_t = 0.19m_0$ (Table 15.1), are occupied.[69] Still, several experiments showed RT Hall mobilities around 1500 cm^2 V^{-1} s^{-1}, which is not higher than the mobility of intrinsic bulk Si. Nelson and coworkers[70] identified the unexpectedly low RT mobilities as an artifact of conventional Hall measurements in the presence of a spurious parallel channel in the doping supply layer. Such a low-mobility channel at the very location of the ionized donors leads to an underestimation of the channel mobility, because in a standard Hall configuration the respective concentrations in the two layers are weighted by their respective mobilities. At low temperatures, the high mobility 2DEG usually dominates even if there is a larger density of carriers in a parasitic, low-mobility channel. At room temperature, however, phonon scattering leads to greatly reduced mobility differences, and, moreover, the higher density of states of the sixfold degenerate parallel channel leads to an increasing population there. These two factors can result in an overall Hall mobility that significantly underestimates the true mobility in the 2DEG. By investigating samples with a systematic variation of the doping concentration in the supply layer, Nelson *et al.* could extrapolate RT mobilities in excess of 2500 cm^2 V^{-1} s^{-1} for the limiting case of mobile carriers being confined to the channel only. These results could be confirmed by employing a variable magnetic-field Hall technique that allows discrimination between the different contributions to the Hall effect.[71] Evidently, the smaller effective mass together with the modified screening behavior of a 2DEG enhance the RT mobility by almost a factor of 2 over that of pure bulk Si.

The mobility enhancement increases to more than a factor of 3 when compared to a long-channel MOSFET with low channel doping. This effect is mainly due to the greatly reduced contribution of interface scattering in the case of a crystalline Si/SiGe interface as compared to the interface between crystalline Si and amorphous SiO_2 in a MOSFET. This conclusion is corroborated by measurements of the Dingle ratio, which is a measure of the dominant interface scattering mechanism. For that purpose the single-particle relaxation time τ_s is deduced from the amplitude evolution of Shubnikov–de Haas (SdH) oscillations at low magnetic fields (i.e. before the onset of spin splitting), and compared with the transport scattering time τ_t associated with the experimental Hall mobility.[66] The Dingle ratio $\alpha = \tau_t/\tau_s$ allows a principal classification of scattering mechanisms, because of the different statistics involved in the definition of the two scattering times. Dingle ratios close to 1 are observed in MOSFETs, indicating predominantly isotropic short-range scattering at interface roughness and interface charges.[72] In contrast, scattering in the channel of MODQW structures is dominated by long-range, small-angle Coulomb scattering at the ionized donors in the doping supply layer. This mechanism enhances τ_t with respect to τ_s, and thus leads to Dingle ratios much larger than 1.[73]

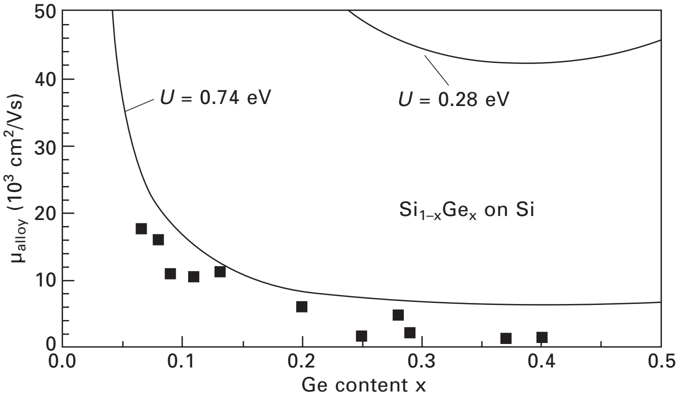
Systematic Dingle experiments were conducted in Ref. 66 on Si/SiGe MODQWs with low-temperature Hall mobilities ranging from 70,000 to 175,000 cm^2 V^{-1} s^{-1}. All samples revealed $\alpha >> 1$, with a maximum value of 26 observed on the sample with the highest mobility. This value is in good agreement with theoretical expectations,[74,17] and is also comparable to results found in III–V MODQWs. Thus, long-range, small-angle scattering is, as expected, the dominating mechanism in good Si/SiGe heterostructures.

15.4.2 Two-dimensional hole gases (2DHG)

Pseudomorphic SiGe quantum wells on unstrained Si

The band offset between a pseudomorphic $Si_{1-x}Ge_x$ layer and an unstrained Si substrate is almost exclusively restricted to the valence band. Hence, a layer sequence utilizing a $Si_{1-x}Ge_x$ channel is the simplest approach toward the implementation of a p-type MODQW structure. Several groups have exploited this concept, and in recent years a substantial amount of transport data has been published. The low-temperature hole mobilities in such structures with compositions limited to $x < 0.5$ because of critical thickness concerns have shown a continuous increase over the years, with the best reported values being close to 20,000 $cm^2 V^{-1} s^{-1}$. On first sight, this appears as an enhancement by a factor of 6 as compared to the very first such structures.[75] However, upon closer inspection it turns out that a significant part of this improvement was correlated with a reduction of the Ge content to values below 15%. The highest mobilities were actually observed at around $x = 10\%$. Although these improved mobilities are certainly an important achievement with respect to a deeper understanding of the materials properties, such low-composition layer sequences have their limitations in the comparably small valence band offset, which restricts the 2D carrier densities to the low 10^{11} cm^{-2} range.

Figure 15.8 shows a compilation of reported low-temperature (<4.2 K)



15.8 Hole mobilities versus Ge content in pseudomorphic $Si_{1-x}Ge_x$ layers. Data points are from Ref. 114. The solid lines are calculated mobilities from Ref. 9 for alloy scattering with different values of the scattering potential *U.*

Hall mobility data for pseudomorphic $Si_{1-x}Ge_x$ channels on Si substrates. Obviously, there is a monotonic decrease of the mobility as x increases. This is in contrast to most of the early theoretical expectations, which predicted a more or less pronounced increase of the hole mobility with composition, caused by the strain-induced reduction of the effective hole mass.[76–78] On the other hand, random alloy scattering, which reflects the local fluctuations of the Coulomb potential, becomes more pronounced as x increases, reaching a maximum at $x = 50\%$. In a virtual crystal approximation the corresponding mobility term can be written as[39,79]

$$\mu_{\text{alloy}} = \frac{16e\hbar^3\Omega_0}{3(m^*_{\text{hh}})^2 b} \cdot \frac{1}{x(1-x)U^2} \tag{15.11}$$

where Ω_0 is the volume per atom, and U is the interaction potential. The parameter b depends on the wave function of the lowest subband, and reads, assuming a simple Fang–Howard wave function and negligible background doping,[63] as

$$b = \left[\frac{12m^*_{\text{hh}}e^2}{\hbar^2\varepsilon\varepsilon_0} \cdot \frac{11}{32}n\right]^{1/3} \tag{15.12}$$

with the 2DHG density n. Because the virtual crystal approximation is based on short-range scatterers, screening has been neglected in eq. 15.11, which leads to the rather weak $n^{-1/3}$ dependence of μ_{alloy} according to eq. 15.12.

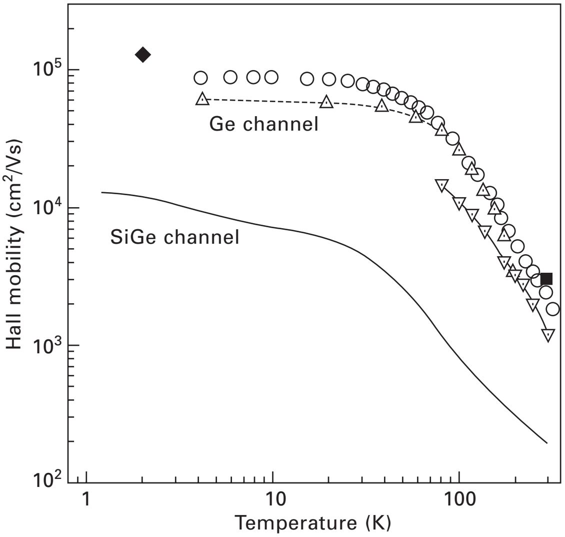
Even in the simple form of eq. 15.11, alloy scattering depends on the effective mass $m_{\text{hh}}*$, on the 2DHG density n, on the scattering potential U, and on the Ge content x. In addition, the effective mass itself is a function of x and n, since strain reduces $m_{\text{hh}}*$ as x increases, whereas the pronounced non-parabolicity of the heavy hole band leads to an increase of $m_{\text{hh}}*$ as n is raised.[80] These effects are coupled, since the higher valence band offsets achievable with increasing x are usually exploited for enhancing n. This way, the strain-induced mass reduction will be partly compensated by non-parabolicity effects. The scattering potential U is given by the change of the valence band offset with composition: $U = d(\Delta E_v)/dx$. The value of U has been discussed controversially in the literature, and a wide range of values between 0.28[21,81] and 0.74[9,16] have been assumed for simulating the influence of alloy scattering. The upper limit of U is given by the unscreened band offset between pure Si and pure Ge, whereas the lower limit is based on fits to field-effect mobilities measured on early heterobipolar transistors. Because of the U^{-2} dependence of μ_{alloy}, the value of U is the most important parameter for judging the importance of alloy scattering.

The expected contribution of alloy scattering for the aforementioned limiting values of U are plotted as solid lines in Fig. 15.8 together with experimental low-temperature mobilities. Both the variations of the effective mass as a function of x and n, and the maximum possible increase of n with x, were accounted for.[9] The curve with $U = 0.74$ eV, which contains no adjustable fitting parameter, gives amazingly well an upper boundary for the experimental data points, closely following the overall mobility decrease as x increases with the characteristic $1/[x(1 - x)]$ fingerprint of alloy scattering in eq. 15.11. Fischetti and Laux came to essentially the same conclusion, with $U = 0.7$ eV resulting from a fit to experimental bulk transport data.[16]

Pseudomorphic Ge quantum wells on Ge-rich metamorphic substrates

The failure to fabricate pseudomorphic $Si_{1-x}Ge_x$ MODQWs with reasonable carrier densities and enhanced mobilities led several groups to the conclusion that pure Ge channels, or at least Ge-rich channels with $x > 70\%$, would be a more promising alternative. In either case, critical thickness considerations in connection with quantum confinement energies do not allow pseudomorphic growth on Si substrates. For an assessment of the ultimate performance some research groups employed Ge substrates,[82,83] but most of the investigations are based on metamorphic substrates with relaxed SiGe buffer layers, similar to the ones employed for the n-channel MODQWs. In their pioneering work Murakami *et al.*[82] used constant composition buffers on a Ge substrate. More recently, thin, constant composition buffers became interesting because of the poor thermal conductivity of $Si_{1-x}Ge_x$. First implementations were reported, which are based on a low-temperature seeding layer to aid the formation of misfit dislocations.[84] However, most of the systematic studies on pure Ge channels have employed the well-tried linearly graded buffers. After some frustrating attempts with constant growth temperatures T_s throughout buffer deposition, down-ramping of T_s with increasing x was introduced to suppress three-dimensional Stranski-Krastanow growth. Also, the strained Ge channel requires low-temperature deposition in order to maintain abrupt interfaces. Especially the upper surface of the channel is prone to 3D growth,[85] which caused several groups to implement inverted modulation doping sequences with the doping supply layer located beneath the channel. Carriers then become located at the lower interface, the roughness of which is more easily controllable. Sawano *et al.* report recently significantly improved Dingle ratios for such an inverted configuration.[86]

Despite the growth complications in the Ge-rich regime, the superior mobility behavior of strained Ge channels is beyond doubt. This is illustrated in Fig. 15.9 which is a compilation of temperature-dependent hole Hall mobilities of Ge-channel MODQWs in comparison with Hall mobilities of one of the best pseudomorphic SiGe-channel structures available.[87] Low-temperature mobilities as high as 120,000 cm^2 V^{-1} s^{-1} were reported for a



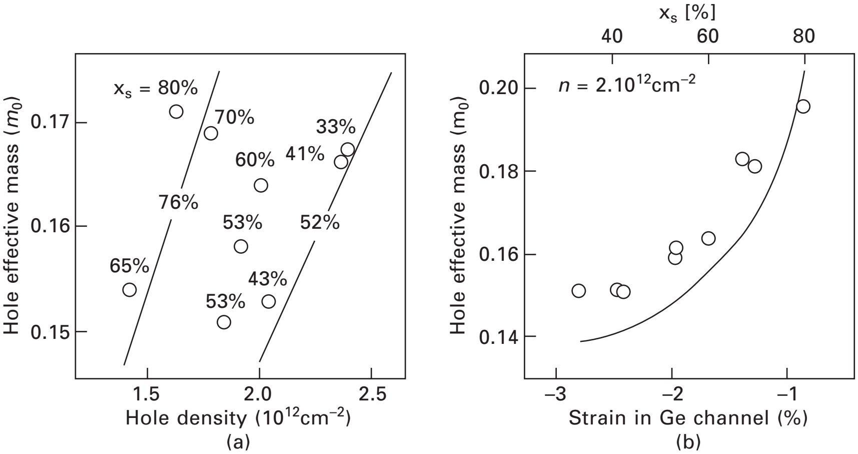
15.9 Temperature variation of the Hall mobilities in Ge-channel MODQWs on relaxed $Si_{1-xs}Ge_{xs}$ buffers, (From Refs 81, 84, 88, 115) and, for comparison, of a pseudomorphic SiGe channel on a Si substrate (From Ref. 86).

sample grown with normal doping sequence by low-energy plasma-enhanced chemical vapor deposition (LEPECVD).[88,89]

A systematic assessment of the mobility-limiting scattering mechanisms in the framework of a linear response theory was reported in Ref. 89 for a series of samples with increasing *n*. The authors found that remote impurity scattering clearly dominates, whereas interface roughness and interface charge scattering play no role. Hence, the best strained Ge channel p-type Si/SiGe MODQWs behave at low temperatures similarly to the strained Si channel n-type MODQW structures. Evidently, interface roughness and interface charges are not relevant in either of these structures. It is therefore not clear why these mechanisms should become dominant in SiGe-channel MODQW structures, as has been proposed.[13,90]

Measurements of the room-temperature Hall mobilities in Ge channel MODQWs have the same problem as their n-channel counterparts, namely an underestimation of the channel mobility due to the development of a parasitic parallel channel at high temperatures. To overcome this measurement artifact, again magnetic-field-dependent Hall measurements have been employed to separate the carrier densities and mobilities in the conducting parallel channels. On this basis, several groups were able to extract the true contribution of the 2DHG in the Ge channels. From their results it can be concluded that room-temperature mobilities between 2400 and 3100 cm^2 V^{-1} s^{-1} can now routinely be achieved in several laboratories.[91–94] This is within range of the predicted room temperature mobility[16] in a strained Ge channel of about 4000 cm^2 V^{-1} s^{-1}.

In the strained Si channel of a Si/SiGe MODQW the electron transport mass is close to the transversal mass $m_t = 0.19m_0$, and, moreover, it is almost independent of the strain in the channel.[10,16,22,95] This is certainly not the case for a strained Ge channel, for which strain-dependent heavy-hole/light-hole splitting and non-parabolicity are expected to have strong influence on the hole transport mass. Sawano *et al.*[86] addressed this problem with a series of Ge-channel MODQWs in which the biaxial compressive strain was varied from –3% to –1% by employing metamorphic substrates with compositions ranging from $0.33 < x < 0.65$. Figure 15.10(a) shows in the upper panel the influence of non-parabolicity, which leads to an increase of the effective hole mass with increasing carrier concentration because of band filling. Figure 15.10(b) shows the transport mass as a function of biaxial strain in the channel at an interpolated carrier concentration of 2×10^{12} cm^{-2}. Within the investigated strain range the effective mass decreases from $m_{\text{hh}}^{*} = 0.19\ m_0$ to 0.15 m_0 with increasing amount of compressive strain. This means that as the heavy hole band is shifted upward and becomes increasingly separated from the light hole bands, it simultaneously increases its curvature, and thus causes a reduced effective mass. These experimental results are in excellent agreement with the predictions of Fischetti and Laux.[16]



15.10 Effective hole mass in compressively strained Ge channels. (a) Influence of non-parabolicity due to band filling with increasing carrier concentration *n* in the channel. Lines are calculated curves. Numbers at the experimental data points refer to the respective composition x_s of the metamorphic substrate. (b) Transport mass as a function of biaxial strain in the channel at an interpolated carrier concentration of 2×10^{12} cm^{-2}. (From Ref. 85)

15.5 Carrier transport in silicon/silicon–germanium (Si/SiGe) devices

15.5.1 Heterobipolar transistors

The first commercial devices utilizing SiGe were heterobipolar transistors (HBT), which were developed in the late 1980s, initially mainly at IBM[96] and Daimler-Benz.[97] The Si/SiGe HBT is the most straightforward device in this material system, because it uses only one additional, pseudomorphic SiGe epilayer for the base of the transistor, but otherwise induces very few changes to the process flow. The purpose of the SiGe base is the introduction of a band offset between the emitter and the base, which efficiently suppresses hole injection from the base into the emitter.[10] The band alignment of the pseudomorphic Si/SiGe heterosystem is ideally suited for npn HBTs, because the band offset occurs almost exclusively in the valence band, whereas no additional barrier is introduced for electron injection from the emitter into the base. The suppressed hole injection into the emitter allows for higher base doping concentrations, and thus both thinner base widths and reduced base sheet resistances. Both parameters lead to an increase of the cutoff frequencies f_T and f_{max}, and thus to intrinsically faster Si-based transistors, which are used either in bipolar circuits or in integrated BiCMOS applications.

Since their first demonstration, the SiGe HBT has become a mature device with ever increasing cutoff frequencies that exceed 300 GHz in the most advanced versions.[10] For more than a decade, SiGe HBTs have belonged to the standard device inventory of most Si device manufacturers and foundries. For further information, the reader is referred to Chapter 19 in this book, and to dedicated review articles on SiGe HBTs.[10,98,99]

15.5.2 Modulation-doped Si/SiGe field effect transistors

One of the motivations for the work on MODQW structures was the prospect of high-speed, low-noise field effect transistors (FET). The concept of mobility enhancement by remote doping was originally introduced by the III–V community,[100] and MODFETs (or HEMTs: high electron mobility transistors) on the base of III–V compound semiconductors are produced in large quantities for low-noise microwave applications. As we have seen in the preceding sections, the Si/SiGe heterosystem requires metamorphic substrates for n-type devices to define an electron channel in a tensile in-plane strained Si layer. Even though pseudomorphic SiGe layers on Si would define a hole channel for p-type devices, we have seen that alloy scattering destroys any improvement introduced by effective mass reduction.[9,16] Therefore, a metamorphic SiGe substrate is also required for p-type devices, which allows the implementation of a Ge-rich SiGe, or, preferentially, a pure Ge channel.

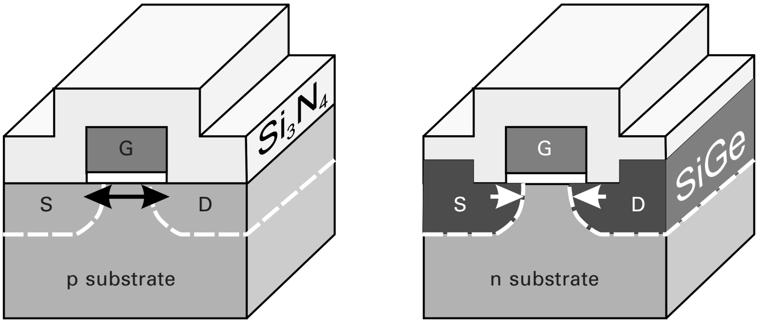
Several research groups have realized and subsequently optimized Si/SiGe MODFETs, most of them controlled by Schottky gates, like their III–V counterparts.[101–104] Over the years a steady improvement of f_T and f_{max} as well as of the transconductances g_m could be demonstrated, for both n- and p-channel MODFETs. The best f_{max} values have reached 200 GHz for channel lengths below 100 nm, and g_m values as high as 500 mS/mm have been reported.[10] While these results are comparable to MOSFETs of the same gate length, the high-frequency noise behavior of the MODFETs is much superior.[105] The excellent noise figures are attributed to the fact that the channel mobility in a MODFET is typically a factor of 10 higher than in a MOSFET with sub-micron gate length.

Despite significant improvements over the years, and low noise characteristics, MODFETs have not yet entered production. The main problems are the incompatibility of the process flow with standard complementary metal oxide semiconductor (CMOS) technology, and the high leakage currents of Schottky gates compared to the MOS (metal-oxide-semiconductor) gates in mainstream Si technology. In addition, SiGe MODFETs compete with III–V HEMTs, which are still leading in performance at costs which are presumably not much higher than those of a potential stand-alone SiGe MODFET technology.

15.5.3 Strained CMOS

While the modulation-doping concept is widely incompatible with standard CMOS technology, CMOS downscaling itself had reached a critical limit by around 2003. By then the reduction of the oxide thickness to just 0.8 nm had led to a drastic increase of the gate leakage current by tunneling. Together with a high overdrive (i.e. the excess gate voltages applied above threshold) to achieve high switching speeds, power dissipation in highly integrated CMOS circuits came close to its very limits.[5] This problem, which is often addressed as ‘the power crisis in the Si chip industry’, could not easily be solved, because high-*k* dielectrics, which would have reduced the gate leakage currents, were at that time not ready for introduction into the process flow. Hence, the only way to reduce the overdrive without further lowering the oxide thickness or the channel length was by increasing the channel mobility.

This path was followed by Intel with the introduction of their 90 nm process in 2003, which employed the first generation of strain engineering in their CMOS process flow.[106] In contrast to the approaches in the preceding sections, which always involved a metamorphic substrate, Intel introduced uniaxial external strain into an otherwise widely unchanged Si MOSFET process. This was done in a different manner for the n- and p-type MOSFETs (Fig. 15.11). The n-channel devices received tensile strain along the [110]



(a) Strained NMOS (b) Strained PMOS

15.11 Schematic view of the stressor techniques first introduced by Intel in 2003 for incorporating uniaxial strain into the Si channels of (a) n-type and (b) p-type MOSFETs. The former received a tensile straining Si_3N_4 cap above the transistor, whereas the latter have their source and drain contacts made of coherently grown SiGe strips (source: Intel).

oriented Si channel by a high-stress SiN film that capped the complete transistor. In the p-type devices compressive strain along the [110] oriented Si channel was introduced by source and drain regions made from highly doped, coherent SiGe. In both cases the stress becomes uniaxial because of the transistor geometry, in which the channel length in the longitudinal direction (between source and drain) is much shorter than the channel width in the transversal direction.

The improvements Intel could achieve with its first-generation strain-engineering process are quite impressive regarding the p-MOSFETs. Uniaxial compressive strain turned out to enhance very efficiently the hole mobility. Finite element calculations revealed that with a Ge concentration of 17% in the SiGe stressors, longitudinal stress on the order of 500 MPa can be induced in a 45nm long channel.[106] This leads to an estimated hole mobility enhancement of 40%, a value that was found to be consistent with measured device data.[107] In contrast, the mobility enhancement achieved by the stressor cap in the n-MOSFETs was only about 10%, because of the much lower stress values achievable with this technique,[106] but also because strain leads only to a higher population of the Δ_2 valleys but not to a change of their transport mass.[16] The lower enhancement of the transport properties in n-channel devices was found to be acceptable, because the electron mobilities in CMOS circuits are intrinsically higher than the hole mobilities.

From a process-flow point of view, Intel's concept of longitudinal strain in the Si channels of standard MOSFETs has several advantages as compared to biaxial strain introduced by metamorphic layers or by active pseudomorphic SiGe layers. First, standard substrates are used. Secondly, current flow is always in a Si layer, i.e. no problems with oxidation, implantation or alloy scattering in a SiGe layers are encountered. Thirdly, n- and p-type devices are optimized separately; and lastly, as channel lengths shrink further, the strain induced by the stressors increases automatically. Therefore, the uniaxial strain concept was certainly the easiest way to bring strain engineering and SiGe into a CMOS environment. In addition, this concept turned out to be scalable: in 2009 Intel introduced its 32 nm platform, which already contained the fourth generation of strain engineering, still essentially based on the original concept.[108] Improvements were achieved by enhancing strain via the recently introduced metal-gate refilling process, and by increasing the Ge content in the $Si_{1-x}Ge_x$ stressors of the p-MOSFETs.[109]

Nevertheless, there are limitations, especially regarding n-type devices. The SiN stressor cap is not very efficient regarding the amount of stress that can be realized in the channel. So far, the intrinsically higher performance of n-type MOSFETs has made the lower performance gain by strain engineering acceptable. In the course of further downscaling, however, more efficient means will be necessary to improve the channel mobilities in n-MOSFETs.

For more information on strained MOSFET devices the reader is referred to Chapters 19 and 20.

15.6 Future trends

Mobility enhancement by strain engineering in the Si/SiGe heterosystem has become an important ingredient of the semiconductor road map,[6] and for several years one of the most important contributions to the fulfillment of Moore's law.[5] Future device developments will certainly contain strain engineering. Several such approaches are being followed at the present time. The introduction of metamorphic $Si_{1-x}Ge_x$ substrates with compositions exceeding 25% is one of the options, which will further enhance n- and p-MOSFET performance.[110–112] The introduction of such substrates will most likely be combined with the large-scale introduction of silicon on insulator (SOI) substrates, which allow for fully depleted devices. Since SOI techniques based on wafer bonding are rapidly advancing, entirely SiGe-free strained Si substrates are likely (SSDOI: strained silicon directly on insulator).[113] Also, different substrate orientations are being investigated, even in a hybrid form, where n- and p-MOSFETs sit on differently oriented parts of a common substrate that is again produced by wafer bonding.[114] In this way, the advantageous mobilities of p-channel devices on (011) substrates could be exploited without affecting the n-MOSFETs, which have their best performance on (001) oriented substrates.

Much further advanced are quantum computing concepts based on the extremely long electron spin coherence times in Si. But even though such concepts employ spins and spin entanglement rather than charge, most of the schemes discussed so far require electron transport for setting the initial stage, for switching on and off the entanglement between electrons, and for the final readout of the result.[115] Coupled quantum dots with zero-dimensional carrier confinement[116] could become the basic building blocks of a future quantum computer, and the unrivaled Si technology will by then reach routinely the lateral dimensions required for the entanglement of many qubits.[117]

15.7 References

1 G. Abstreiter, H. Brugger, T. Wolf, H. Jorke, H.-J. Herzog, *Phys. Rev. Lett.* 54, 2441 (1985)
2 A. M. Tyryshkin, S. A. Lyon, A. V. Astashkin, A. M. Raitsimring, *Phys. Rev. B* 68, 193207 (2003)
3 A. M. Tyryshkin, S. A. Lyon, W. Jantsch, F. Schäffler, *Phys. Rev. Lett.* 94, 126802 (2005)
4 G. E. Moore, *Electronics Magazine* 38, 4 (1965)
5 M. M. Ieong, B. Doris, J. Kedzierski, K. Rim, M. Yang, *Science* 306, 2057 (2004)
6 See the actual version of the biannual International Technology Roadmap for Semiconductors at http://public.itrs.net/reports.html
7 S. Natarajan *et al.*, *IEEE Techn. Digest IEDM* 2008, 1–3
8 P. Packan *et al.*, *IEEE Techn. Digest IEDM* 2009
9 F. Schäffler, *Semicond. Sci. Technol.* 12, 1515 (1997)
10 D. J. Paul, *Semicond. Sci. Technol.* 19, R75 (2004)
11 E. Kasper (ed.), *Properties of Strained and Relaxed Silicon Germanium*, EMIS Datareviews Series, No. 12, INSPEC, London (1995)
12 F. Schäffler, chapter 6: silicon–germanium, in *Properties of Advanced Semiconductor Materials*, M. E. Levinshtein, S. L. Rumyantsev, M. S. Shur (eds), John Wiley & Sons, New York (2001), pp 149–187
13 T. E. Whall, *J. Cryst. Growth* 157, 353 (1995)
14 L. E. Kay, T.-W. Tang, *J. Appl. Phys.* 70, 1483 (1991)
15 D. Monroe, Y. H. Xie, E. A. Fitzgerald, P. J. Silverman, G. P. Watson, *J. Vac. Sci. Technol.* B 11, 1731 (1993)
16 M. V. Fischetti, S. E. Laux, *J. Appl. Phys.* 80, 2234 (1996)
17 F. Stern, S. E. Laux, *Appl. Phys. Lett.* 61, 1110 (1992)
18 D. B. M. Klaassen, *Solid State Electronics* 35, 953 (1992)
19 S. Krishnamurthy, A. Sher, A.-B. Chen, *Appl. Phys. Lett.* 47, 160 (1985)
20 E. Braunstein, *Phys. Rev.* 130, 869 (1963)
21 T. Manku, A. Nathan, *IEEE Electron Device Lett.* 12, 704 (1991)
22 M. Rieger, P. Vogl, *Phys. Rev. B* 48, 14276 (1993)
23 C. Van de Walle, R. M. Martin, *Phys. Rev. B* 34, 5621 (1986)
24 L. Colombo, R. Resta, S. Baroni, *Phys. Rev. B* 44, 5572 (1991)
25 D. Rideau, M. Feraille, L. Campolini, M. Minondo, C. Tavernier, H. Jaouen, *Phys. Rev. B* 74, 195208 (2006)
26 L. Hedin, *Phys. Rev.* 139, A796 (1965)
27 C. Penn, F. Schäffler, G. Bauer, S. Glutsch, *Phys. Rev. B* 59, 13314 (1999)
28 H. Jorke, H.-J. Herzog, *Proc. 1st Int. Symp. on Silicon MBE*, vol 85–7, J. C. Bean (ed.) Electrochemical Society, Pennington, NJ (1985), p. 194

29 T. Okamoto, K. Hosoya, S. Kawaji, A. Yagi, A. Yutani, Y. Shiraki, *Physica E* 6, 260 (2000)

30 K. Toyama, T. Nishioka, K. Sawano, Y. Shiraki, T. Okamoto, *Phys. Rev. Lett.* 101, 016805 (2008)

31 F. Schäffler, D. Többen, H.-J. Herzog, G. Abstreiter, B. Holländer, *Semicond. Sci. Technol.* 7, 260 (1992)

32 K. Ismail, F. K. LeGoues, K. L. Saenger, M. Arafa, J. O. Chu, P. M. Mooney, B. S. Meyerson, *Phys. Rev. Lett.* 73, 3447 (1994)

33 K. Lai, W. Pan, D. C. Tsui, S. Lyon, M. Mühlberger, F. Schäffler, *Phys. Rev. Lett.* 93, 156805 (2004)

34 J. Matsunami, M. Ooya, T. Okamoto, *Phys. Rev. Lett.* 97, 066602 (2006)

35 Z. Wilamowski, N. Sandersfeld, W. Jantsch, D. Többen, F. Schäffler, *Phys. Rev. Lett.* 87, 026401 (2001)

36 K. Lai, W. Pan, D. C. Tsui, S. Lyon, M. Mühlberger, F. Schäffler, *Phys. Rev. B* 75, 033314 (2007)

37 W. Schmelz, D. Pachinger, F. Schäffler, unpublished

38 T.-M. Lu, D. C. Tsu, D. Pachinger, F. Schäffler, unpublished

39 G. Bastard, in *Wave Mechanics Applied to Semiconductor Heterostructures*, Les Editions de Physique, Paris (1990)

40 T. Okamoto, M. Ooya, K. Hosoya, S. Kawaji, *Phys. Rev. B* 69, 041202 (2004)

41 K. von Klitzing, G. Dorda, M. Pepper, *Phys. Rev. Lett.* 45, 494 (1989)

42 D. C. Tsui, H. L. Stormer, A. C. Gossard, *Phys. Rev. Lett.* 48, 1559 (1982)

43 R. B. Laughlin, *Phys. Rev. Lett.* 50, 1395 (1983)

44 B. Koiller, X. Hu, S. Das Sarma, *Phys. Rev. Lett.* 88, 027903 (2002)

45 A. L. Saraiva, M. J. Calderon, X. Hu, S. Das Sarma, B. Koiller, *Phys. Rev. B* 80, 081305R (2009)

46 D. Monroe, Y. H. Xie, E. A. Fitzgerald, P. J. Silverman, *Phys. Rev. B* 46, 7935 (1992)

47 S. F. Nelson, K. Ismail, J. O. Chu, B. S. Meyerson, *Appl. Phys. Lett.* 63, 367 (1993)

48 K. Ismail, M. Arafa, K. L. Saenger, J. O. Chu, B. S. Meyerson, *Appl. Phys. Lett.* 66, 1077 (1995)

49 P. Weitz, R. J. Haug, K. von Klitzing, F. Schäffler, *Surf. Sci.* 361/362, 542 (1996)

50 R. B. Dunford, N. Griffin, D. J. Paul, M. Pepper, D. J. Robbins, A. C. Churchill, W. Y. Leong, *Thin Solid Films* 369, 316 (2000)

51 K. Lai, W. Pan, D. C. Tsui, S. Lyon, M. Mühlberger, F. Schäffler, *Phys. Rev. Lett.* 93, 156805 (2004)

52 For a review, see, e.g., O. Heinonen (ed.), *Composite Fermions: A Unified View of the Quantum Hall Regime*, World Scientific, Singapore (1998)

53 V. S. Khrapai, A. A. Shashkin, V. T. Dolgopolov, *Phys. Rev. B* 67, 113305 (2003)

54 R. Weitz, R. Haug, K. von Klitzing, F. Schäffler, *Surf. Sci.* 361–362, 562 (1996)

55 K. Lai, W. Pan, D. C. Tsui, S. Lyon, M. Mühlberger, F. Schäffler, *Phys. Rev. Lett.* 96, 076805 (2006)

56 K. Lai, T. M. Lu, W. Pan, D. C. Tsui, S. Lyon, J. Liu, Y. H. Xie, M. Mühlberger, F. Schäffler, *Phys. Rev. B* 73, 161301(R) (2006)

57 T. Jungwirth, S. P. Shukla, L. Smrcka, M. Shayegan, A. H. MacDonald, *Phys. Rev. Lett.* 81, 2328 (1998)

58 K. Lai, W. Pan, D. C. Tsui, S. Lyon, M. Mühlberger, F. Schäffler, *Physica E* 34, 176 (2006)

59 K. Toyama, T. Nishioka, K. Sawano, Y. Shiraki, T. Okamoto, *Phys. Rev. Lett.* 101, 016805 (2008)

60 U. Zeitler, H. W. Schumacher, A. G. M. Jansen, R. J. Haug, *Phys. Rev. Lett.* 86, 866 (2001)

61 S. Goswami, K. A. Slinker, M. Friesen, L. M. McGuire, J. L. Truitt, C. Tahan, L. J. Klein, J. O. Chu, P. M. Mooney, D. W. Van der Weide, R. Joynt, S. N. Coppersmith, M. A. Eriksson, *Nature Physics* 3, 41 (2007)

62 L. M. McGuire, M. Friesen, K. A. Slinker, S. N. Coppersmith, M. A. Eriksson, *N. J. Phys.* 12, 033039 (2010)

63 T. Ando, A. B. Fowler, F. Stern, *Rev. Mod. Phys.* 54, 437 (1982)

64 T. B. Boykin, G. Klimeck, M. A. Eriksson, M. Friesen, S. N. Coppersmith, *Appl. Phys. Lett.* 84, 115 (2004)

65 D. Többen, D. A.Wharam, G. Abstreiter, J. P. Kotthaus, F. Schäffler, *Semicond. Sci. Technol.* 10, 711

66 D. Többen, F. Schäffler, A. Zrenner, G. Abstreiter, *Phys. Rev. B* 46, 4344 (1992)

67 G. Scappucci, L. Di Gaspare, E. Giovine, A. Notargiacomo, R. Leoni, E. Evangelisti, *Phys. Rev. B* 74, 035321 (2006)

68 M. A. Wilde, M. Rhode, C. Heyn, D. Heitmann, D. Grundler, U. Zeitler, F. Schäffler, R. J. Haug, *Phys. Rev. B* 72, 165429 (2005)

69 S. Q. Murphy, Z. Schlesinger, S. F. Nelson, J. O. Chu, B. S. Meyerson, *Appl. Phys. Lett.* 63, 222 (1993)

70 S. F. Nelson, K. Ismail, J. O. Chu, B. S. Meyerson, *Appl. Phys. Lett.* 63, 367 (1993)

71 H. Brugger, H. Koser, *III-Vs Review* 8, 41 (1995)

72 F. F. Fang, A. B. Fowler, A. Hartstein, *Phys. Rev. B* 15, 2127 (1977)

73 M. A. Paalanen, D. C. Tsui, J. C. M. Hwang, *Phys. Rev. Lett.* 51, 2226 (1983)

74 A. Gold, *Phys. Rev. B* 38, 10798 (1988)

75 R. People, J. C. Bean, D. V. Lang, A. M. Sergent, H. L. Störmer, K. W. Wecht, R. T. Lynch, K. Baldwin, *Appl. Phys. Lett.* 45, 1231 (1984)

76 J. M. Hinckley, J. Singh, *Phys. Rev. B* 41, 2912 (1990)

77 T. Manku, A. Nathan, *IEEE Electron Device Lett.* 12, 704 (1991)

78 T. Manku, A. Nathan, *J. Appl. Phys.* 69, 8414 (1991)

79 T. Ando, *J. Phys. Soc. Jpn* 51, 3900 (1982)

80 T. E.Whall, A. D. Plews, N. L. Mattey, P. J. Phillips, U. Ekenberg, *Appl. Phys. Lett.* 66, 2724 (1995)

81 J. M. Hinckley, J. Singh, *Phys. Rev. B* 41, 2912 (1990)

82 E. Murakami, K. Nakagawa, A. Nishida, M. Miyao, *IEEE Electron. Device Lett.* 12, 71 (1991)

83 J. F. Nützel, C. M. Engelhardt, R. Wiesner, D. Többen, M. Holzmann, G. Abstreiter, *J. Cryst. Growth* 150, 1011 (1995)

84 M. Myronow, D. R. Leadley, Y. Shiraki, *Appl. Phys. Lett.* 94, 092108 (2009)

85 Y. H. Xie, D. Monroe, E. A. Fitzgerald, P. J. Silverman, F. A. Thiel, G. P. Watson, *Appl. Phys. Lett.* 63, 2263 (1993)

86 K. Sawano, K. Toyama, R. Masutomi, T. Okamoto, N. Usami, K. Arimoto, K. Nakagawa, Y. Shiraki, *Appl. Phys. Lett.* 95, 122109 (2009)

87 D. J. Paul, N. Griffin, D. D. Arnone, M. Pepper, C. J. Emelius, P. J. Philips, T. E. Whall, *Appl. Phys. Lett.* 69, 2704 (1996)
88 H. von Känel, M. Kummer, G. Isella, E. Müller, T. Hackbarth, *Appl. Phys. Lett.* 80, 2922 (2002)
89 B. Rössner, D. Chrastina, G. Isella, H. von Känel, *Appl. Phys. Lett.* 84, 3058 (2004)
90 J. C. Brighten, I. D. Hawkins, A. R. Peaker, R. A. Kubiak, E. H. C. Parker, T. E. Whall, *Semicond. Sci. Technol.* 8, 1487 (1993)
91 G. Höck, M. Glück, T. Hackbarth, H. J. Herzog, E. Kohn, *Thin Solid Films* 336, 141 (1998)
92 S. Madhavi, V. Venkataraman, Y. H. Xie, *J. Appl. Phys.* 89, 2497 (2001)
93 M. Myronov, T. Irisawa, O. A. Mironov, S. Kohn, Y. Shiraki, T. E. Whall, E. H. C. Parker, *Appl. Phys. Lett.* 80, 3117 (2002)
94 M. Myronov, K. Sawano, Y. Shiraki, T. Mouri, K. M. Itoh, *Appl. Phys. Lett.* 91, 082108 (2007)
95 N. Griffin, D. D. Arnone, D. J. Paul, M. Pepper, D. J. Robbins, A. C. Churchill, J. M. Fernández, *J. Vac. Sci. Technol.* 16, 1655 (1998)
96 G. L. Patton, S. S. Subramanian, S. L. Delage, S. Tiwari, J. M. C. Stork, *IEEE Electron Device Lett.* 9, 165 (1988)
97 A. Schüppen, U. Erben, A. Gruhle, H. Kibbel, H. Schumacher, U. König, *IEEE IEDM Techn. Digest*, 743–746 (1995)
98 S. C. Jain, S. Decoutere, M. Willander, H. E. Maes, *Semicond. Sci. Technol.* 16, R67 (2001)
99 B. S. Meyerson, *IBM J. Res. Dev.* 44, 391 (2000)
100 R. Dingle, H. Stormer, A. C. Gossard, W. Wiegmann, *Appl. Phys. Lett.* 33, 665 (1978)
101 K. Ismail, S. Rishton, J. O. Chu, K. Chan, B. S. Meyerson, *IEEE Electron Device Lett.* 14, 348 (1993)
102 G. Höck, T. Hackbarth, U. Erben, E. Kohn, U. König, *Electron. Lett.* 34, 1888 (1998)
103 A. Kasamatsu, K. Kasai, K. Hikosaka, T. Matsui, T. Mimura, *Appl. Surf. Sci.* 224, 382 (2004)
104 S. J. Koester, K. L. Saenger, J. O. Chu, Q. C. Quyang, J. A. Ott, M. J. Rooks, D. F. Canaperi, J. A. Tornello, C. V. Jahns, S. E. Steen, *Electron. Lett.* 39, 1684 (2003)
105 F. Aniel, M. Enciso-Aguilar, L. Giguerre, P. Crozat, R. Adde, T. Mack, U. Seiler, T. Hackbarth, H. J. Herzog, U. König, B. Raynor, *Solid-State Electron.* 47, 283 (2003)
106 S. E. Thompson *et al.*, *IEEE Trans. Electron Devices* 51,1790 (2004)
107 S. E. Thomson *et al.*, *IEEE Electron Device Lett.* 25, 191 (2004)
108 C.-H. Jan *et al.*, *IEEE Techn. Digest IEDM* 2009
109 C. Auth, *IEEE CICC* 2008, 379
110 K. Rim, J. L. Hoyt, J. F. Gibbons, *IEEE Trans. Electron Device* 47, 1406 (2000)
111 N. Sugii, D. Hisamoto, K. Washio, N. Yokoyama, S. Kimura, *IEEE Trans. Electron Device* 49, 2237 (2002)
112 L. Huang *et al.*, *IEEE Electron Device Lett.* 49, 1566 (2002)
113 K. Rim *et al.*, *IEEE Techn. Digest IEDM* 2003, 49
114 M. Yang *et al.*, *IEEE Techn. Digest IEDM* 2003, 453

115 R. Vrijen, E. Yablonovitch, K. Wang, H. W. Jiang, A. Balandin, V. Roychowdhury, T. Mor, D. DiVincenzo, *Phys. Rev. A* 62, 123306 (2000)

116 N. Shaji, C. B. Simmons, M. Thalakalum, L. J. Klein, H. Quin, H. Luo, D. E. Saavage, M. G. Lagally, A. J. Rimberg, R. Joynt, M. Friesen, R. H. Blick, S. N. Coppersmith, M. A. Eriksson, *Nature Physics* 4, 540 (2008)

117 M. Friesen, R. Joynt, M. A. Eriksson, *Phys. Rev. B* 67, 121301 (2003)

16

Microcavities and quantum cascade laser structures based on silicon–germanium (SiGe) nanostructures

J. XIA and Y. SHIRAKI, Tokyo City University, Japan and
J. YU, Chinese Academy of Sciences, P.R. China

Abstract: This chapter discusses two types of Si-based light-emitting devices based on Ge quantum dots in optical microcavities and SiGe quantum cascade (QC) structures. After reviewing various solutions for Si-based light-emitting devices, the chapter describes the method to enhance light emission from Ge dots through embedding them into optical microcavities. It then reviews SiGe quantum cascade laser (QCL) structures on issues of material growth, electroluminescence from SiGe QC structures, n-type SiGe QC structures, and waveguides for SiGe QCLs.

Key words: silicon–germanium (SiGe), Si-based light-emitting devices, microcavity, germanium (Ge) quantum dots, quantum cascade laser (QCL).

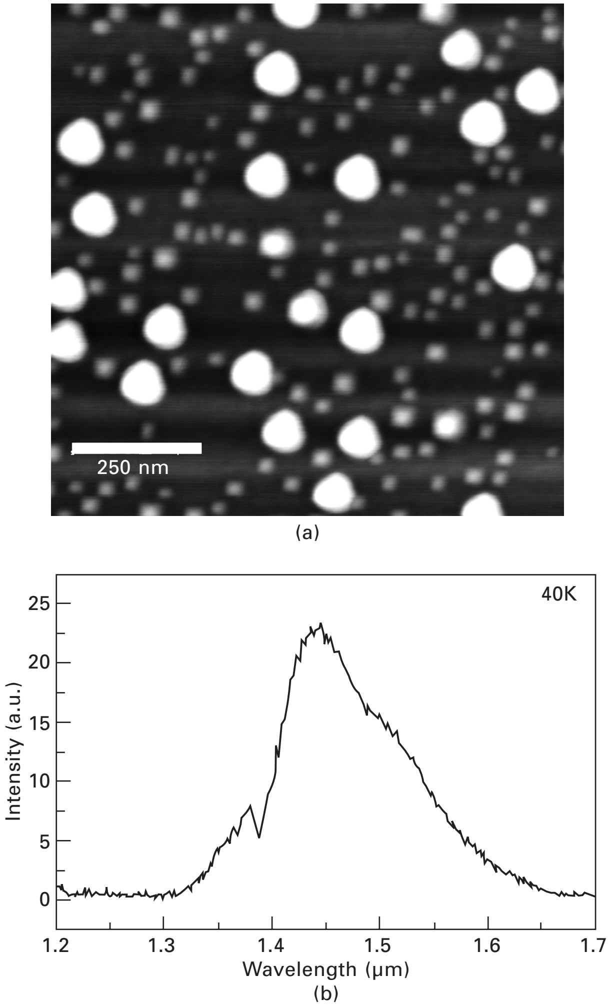
16.1 Introduction

The demand for computational power and data transmission bandwidth is becoming higher and higher due to the rapid development of the information society. The size of transistors is getting smaller and smaller, and current electronic integrated circuits on the Si-platform are facing the limitations of fabrication difficulties, thermal dissipation, and bandwidth. In order to fulfill the growing demand, scientists are performing numerous efforts to develop the next generation of integrated circuits on the Si platform. One of the solutions is optoelectronic hybrid integration on Si which has the advantages of ultra-large bandwidth and low power consumption. The emerging view favors integrating different optical components, such as Si-based light source, optical waveguide circuits, modulator, detector, and electronic integrated circuit for control on a single Si wafer. Therefore, complementary metal oxide semiconductor (CMOS) compatible Si-based light emitting devices are inevitable for optoelectronic integration on Si. Unfortunately, Si is an indirect bandgap material, and it is not a good light-emitting material. In order to develop Si as a light-emitting material, various proposals, such as SiGe quantum wells and dots [1–6], Si/SiGe quantum-cascade structures [7, 8], Er doping [9], Si nanocrystals [10], and defects engineering [11] have been proposed for creating light-emitting centers in Si and Si-based materials. Among these methods, Si-based light emitting devices based on Ge quantum dots and SiGe cascade structures have been attracting much attention since they are fully CMOS-compatible.

16.2 Germanium (Ge) dots microcavity photonic devices

Ge self-assembled quantum dots attract much attention due to their potential as a Si-based light source. Since Ge belongs to the group IV materials and is already being used in the integrated circuit industry, Ge self-assembled quantum dots-based light emitting devices are fully compatible with current CMOS technology. Ge self-assembled quantum dots can be easily grown using molecular beam epitaxy (MBE) in the Stranski–Krastanov (S–K) mode. In Ge quantum dots, the radiative recombination probability is enhanced due to the space trapping of excited carriers by the type-II band structure [3–6]. Figure 16.1(a) shows an atomic force microscope (AFM) image of Ge self-assembled quantum dots. As seen in the figure, two kinds of quantum dots, pyramids and domes, are formed during the growth. The overall dot density is around $1.6 \times 10^{10}/\text{cm}^2$. Figure 16.1(b) shows the photoluminescence (PL) spectrum of the Ge quantum dots measured at 40 K. A wide PL peak covers the wavelength range from 1.3 to 1.6 μm, for which Si is transparent. This is very important for the future integration of silicon-based light sources and optical waveguide circuits on a single Si chip. The wide dispersion of the quantum dot size leads to the broadening of the photoluminescence spectrum. Si/Ge materials are more stable compared with porous and amorphous silicon, silicon nanocrystals and engineered defects in Si. Moreover, it is easy to carry out current injection based on the SiGe platform, which is important for fabricating practicable devices. However, the main drawbacks of light emission from Ge self-assembled dots are (i) lack of spectral purity; (ii) lack of directionality; and (iii) low efficiency.

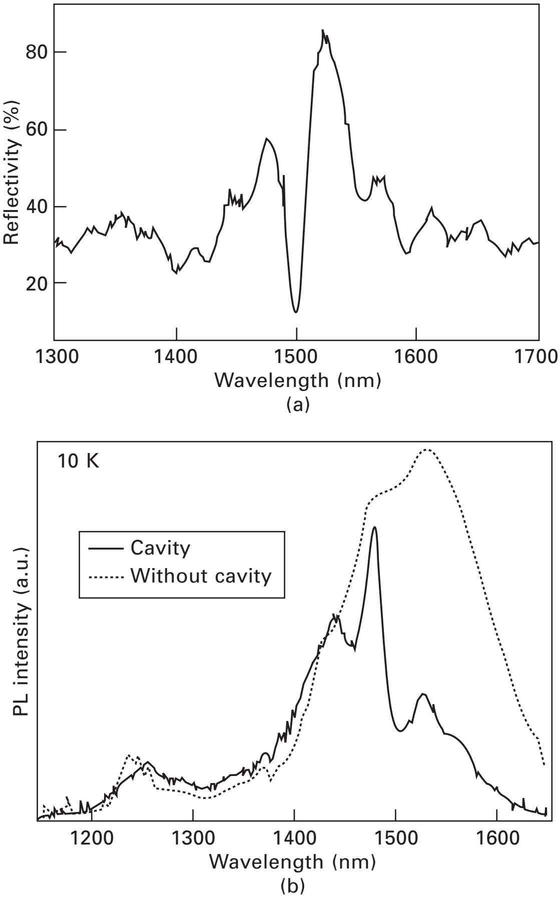
An optical microcavity, which is currently one of the hottest research topics, can solve these problems in principle [12, 13]. Optical microcavities modify the light emission from Ge dots embedded inside due to the optical resonance. The spontaneous emission rates of Ge dots are modified by the electromagnetic field with the cavity. Only the emission at discrete wavelengths, corresponding to the cavity modes, is selected and enhanced, and the other emission is suppressed due to the Purcell effect [12]. The enhancement is proportional to the quality factor over the mode volume of the cavity modes. Different microcavities have been used to modulate, enhance, and control the light emission from Ge self-assembled quantum dots.



16.1 (a) AFM image of the Ge self-assembled quantum dots grown at 600°C by gas-source MBE. (b) PL spectrum of the Ge self-assembled quantum dots at 40 K.

16.2.1 Ge dots in distributed Bragg reflector (DBR) cavities

Kawaguchi and his colleagues embedded Ge self-assembled quantum dots into planar microcavities formed by SiGe/Si DBRs [14–17]. The DBRs were formed by growing strain-balanced $Si_{0.73}Ge_{0.27}$/Si pairs on a relaxed $Si_{0.89}Si_{0.11}$ buffer layer on graded buffer on Si substrate using gas-source molecular beam epitaxy (GS-MBE). A one-wavelength-thick $Si_{0.89}Si_{0.11}$ with



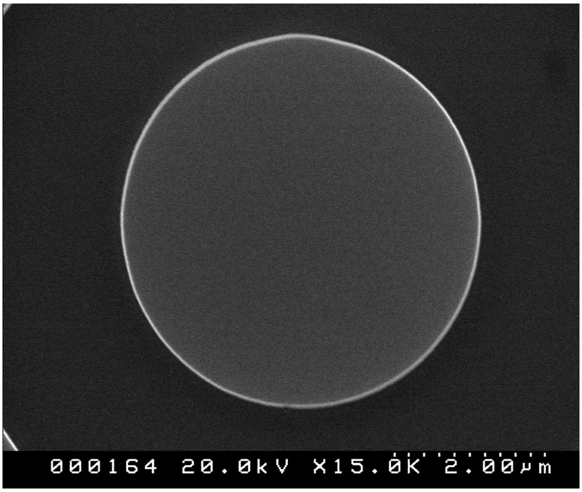
16.2 (a) Reflectivity spectrum of a planar cavity with Ge dots formed by SiGe/Si DBRs. (b) PL spectrum at 10 K from Ge quantum dots in the cavity formed by SiGe/Si DBRs. The dashed line shows the PL spectrum of a reference sample without cavity. Adopted from Ref. 15.

Ge dots was embedded into two DBRs. Figure 16.2(a) shows the reflectivity of the cavity structure at room temperature and Fig. 16.2(b) shows the PL spectrum of the cavity sample together with the PL spectrum of a reference sample without cavity at 10 K. The dip in the reflectivity spectrum indicates the resonant wavelength of the cavity. The shape of the PL spectrum from the cavity sample is completely different from that of the spectrum from the reference sample, showing the function of the cavity. The highest peak is assumed to be the resonant peak of the cavity. Angular dependence of the PL intensity was also observed, giving further evidence for the optical resonance induced by the cavity structure.

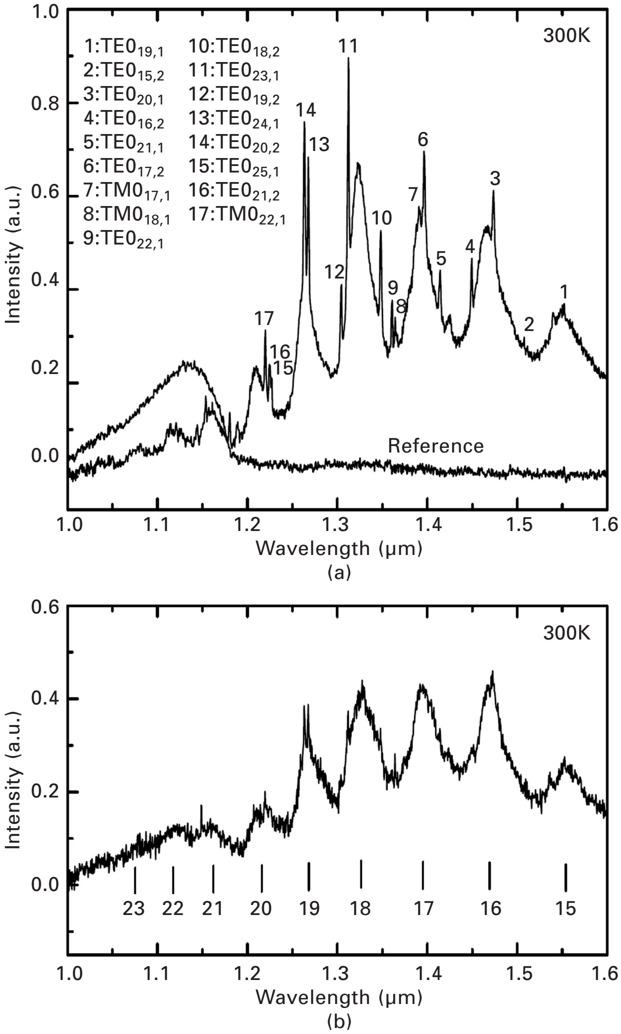
16.2.2 Microdisk and microring resonators with Ge dots

Compared with vertical cavities formed by SiGe/Si DBRs, microdisk and microring [18] resonators are simple and easy to fabricate. Strong and sharp resonant PL has been observed from Ge self-assembled quantum dots in both microdisk and microring resonators [19, 20].

It is well known that microdisk resonators support whispering-gallery modes (WGMs) which are confined close to the edge [18]. Figure 16.3 shows a scanning electron microscope (SEM) image of a fabricated microdisk [19]. In the disk, three layers of Ge dots are embedded inside. Figure 16.4 shows the measured room-temperature μ-PL spectrum of a microdisk with 4-μm diameter when the pumping laser spot was located at different locations of the disk. To activate the WGMs, the pumping laser spot is located close to the edge of the disk for Fig. 16.4(a). A strong PL signal is observed in the range from 1.1 to 1.6 μm. For comparison, the bottom reference PL spectrum is recorded in the unprocessed region on the chip under the same condition. It is a typical room-temperature PL spectrum for silicon and there is no obvious signal in the emission range of Ge dots. Compared with the reference, the PL intensity from Ge quantum dots is significantly enhanced in the microdisk resonator. This enhancement may be attributed to two factors: the Purcell effect [12] and enhancement of extraction efficiency due to the light scattering at the disk edge. Two kinds of resonant peaks, sharp peaks and broad peaks, exist in the spectrum shown in Fig. 16.4(a). The sharp resonant peaks are superimposed on the lower broad peaks. The quality factors of the sharp peaks are in the range from 600 to 800, and the



16.3 SEM image of a fabricated microdisk resonator with Ge quantum dots.

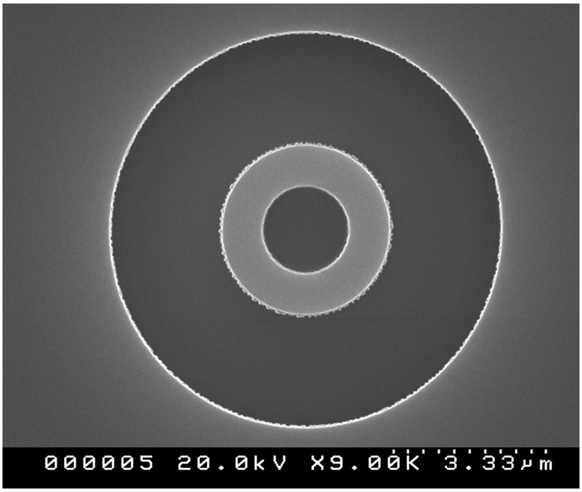


16.4 Room-temperature μ-PL spectra of a 4-μm microdisk with Ge dots. (a) The pumping laser spot is located at the disk edge. The reference spectrum at the bottom is recorded in PhC pattern-free region. (b) The pumping laser spot is located at the disk center.

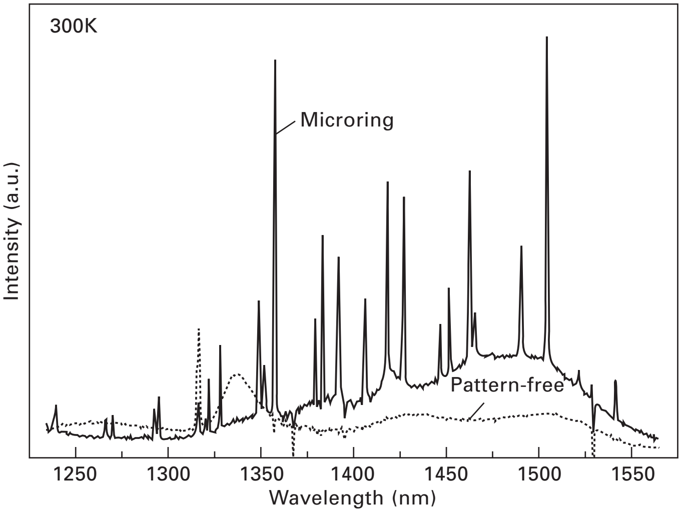
quality factors of the broad peaks are much lower. It is proved that the sharp resonant peaks correspond to the WGMs supported by the disk because they can be activated only when the pumping spot is located at the edge of disk. When the laser spot is located at the disk center, the recorded μ-PL spectrum is shown in Fig. 16.4(b). As can be seen in the figure, the sharp resonant peaks disappear, and only the broad resonant peaks remain. This is reasonable, since the pumping at the center of the disk cannot activate the WGMs because they are confined to the edge of the disk. All the sharp resonant peaks in the range from 1.2 to 1.6 μm are identified by calculations [10, 18] and listed in Fig. 16.4(a). As shown in Fig. 16.4, the broad peaks can be activated at both the edge and the center of the disk. We attribute these broad peaks to Fabry–Pérot (F–P) like modes, which are different from WGM and can be activated everywhere inside the disk. The F–P like modes correspond to another means of resonance: light propagates along the axis of the disk and is reflected by the opposite edges of the disk. The resonant wavelengths for TE-polarized F–P modes are computed and shown by the short vertical lines in Fig. 16.4(b). The numbers below the short lines show the order of the corresponding F–P modes. Because a large amount of optical power escapes from the cavity when light is reflected by the edges of the disk, the quality factors of the F–P modes are much smaller than those of WGMs, as shown in the figure.

The F–P modes in the microdisk are not useful due to their low quality factors. These modes consume part of excited carriers, which leads to a decrease of the intensity of wanted WGMs.

The structure of a microring is similar to that of a microdisk, but the central area is etched to form a ring-shaped waveguide. Since the light propagation along the axis of the disk is terminated by the central air hole, F–P modes are not supported by microrings. The light can only propagate along the ring waveguide, which leads to the WGMs. Figure 16.5 shows the SEM image of a 3-μm microring resonator with Ge dots [20]. The width of the waveguide is 0.7 μm. Ten layers of Ge self-assembled quantum dots were grown on a SOI wafer with a 2-μm buried dioxide (BOX). Figure 16.6 shows the room-temperature μ-PL spectrum of the 3-μm microring resonator. The upper line



16.5 SEM image of a fabricated microring resonator with Ge dots.

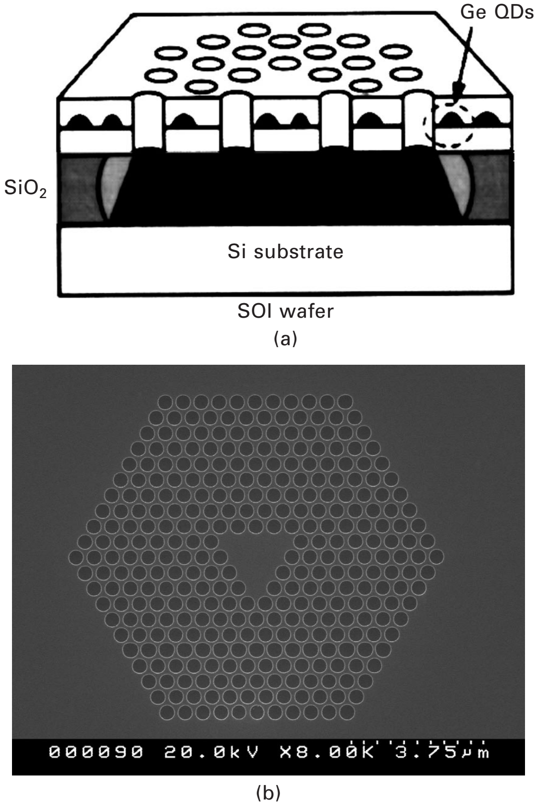


16.6 Room-temperature μ-PL spectra of a 3-μm microring resonator with Ge quantum dots. The top line is recorded in the microring. The bottom line is recorded in the patter–free region under the same conditions.

represents the μ-PL of the microring, while the lower line represents the μ-PL spectrum in the unpatterned region on the same sample under the same configuration. Multiple sharp resonant peaks, corresponding to the WGMs, are observed at room temperature. The quality factors of these resonant peaks are in the range 1800–2200, which are much higher than the value of the photonic crystal (PhC) and microdisk resonators fabricated on SOI wafers with a 400-nm BOX. A highest quality factor around 3000, corresponding to a 0.5-nm line width, has been achieved for a 4-μm microring resonator with a 1-μm waveguide width. The increase of the quality factor is mainly attributed to the increase of the BOX thickness of the SOI wafer used in the microring fabrication, which leads to the dramatic decrease of coupling loss to the Si substrate. Compared with the unpatterned region, the PL intensity is significantly enhanced.

16.2.3 Photonic crystal (PhC) microcavities

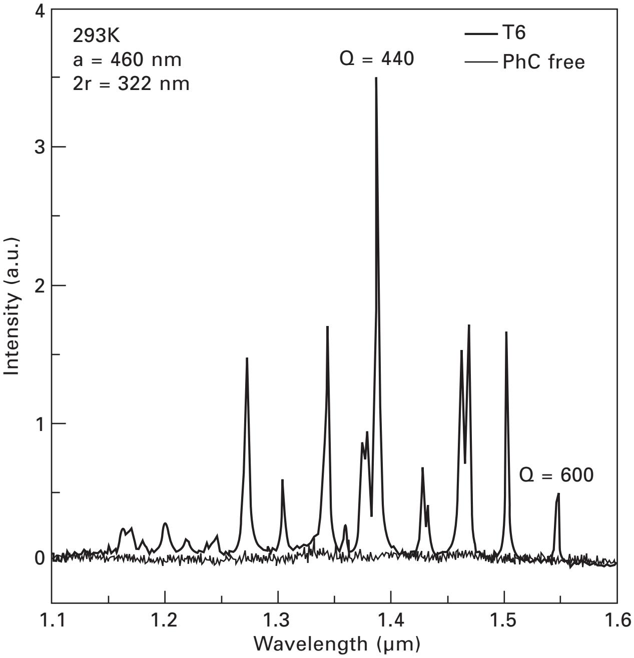
PhC microcavities provide more flexibility in the design of small microcavities with higher performances. For example, the wavelength of resonant peaks can be controlled by tuning the lattice constant of the PhC structure. Combination of Ge self-assembled quantum dots with PhC cavities was proven to be a possible direction for a Si-based light source [20–26]. Figure 16.7(a) shows the schematic structure of the PhC microcavities reported in references [21]



16.7 (a) Schematic structure of the PhC microcavity with Ge self-assembled quantum dots. (b) SEM image of the fabricated T6 cavity.

and [22]. Three layers of Ge self-assembled quantum dots were grown on a SOI wafer with a 400-nm BOX as internal light emitters. The microcavity is formed by introducing defects into a two-dimensional PhC lattice. The PhC lattice is designed to have an optical bandgap covering the emission range of Ge dots from 1.3 to 1.6 μm. Due to the optical bandgap, the surrounding PhC structure provides the in-plane optical confinement, and the optical confinement in the normal direction is provided by total internal reflection (TIR) at the interface between silicon and air. Therefore, three-dimensional optical confinement is achieved by this structure. Figure 16.7(b) shows a SEM image of a fabricated T6 cavity. The triangular-shaped cavity is formed by removing the central 10 air holes. Since the length of the cavity edge is six periods, we denote it as T6.

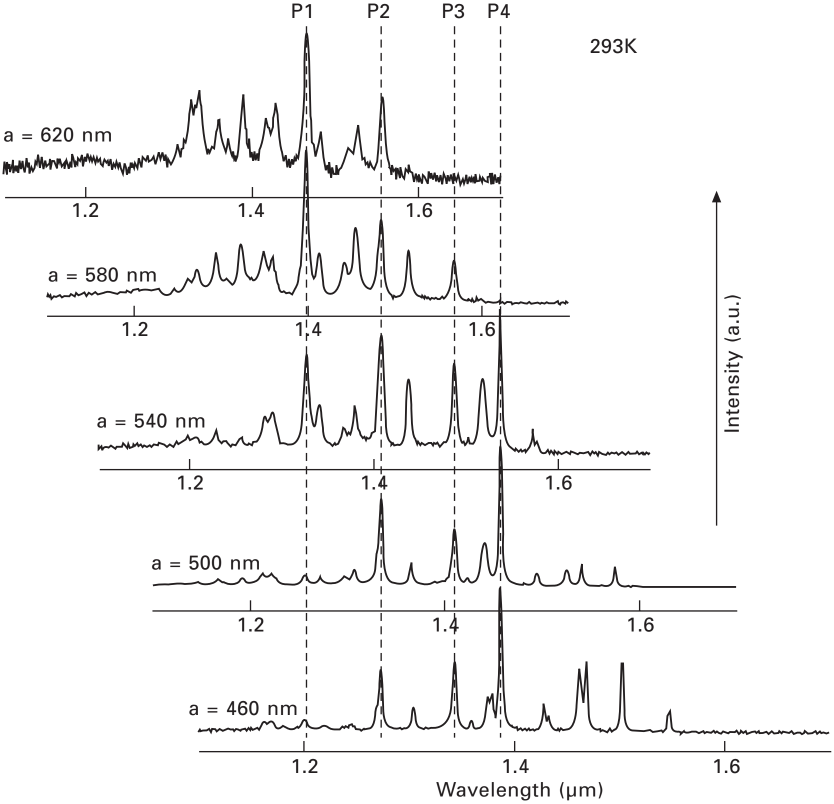
Figure 16.8 shows the room-temperature μ-PL spectrum and reference spectrum from T6 cavity. The PhC lattice constant and the diameter of the air holes are 460 nm and 322 nm respectively. Strong resonant room-temperature light emission is observed. Multiple resonant luminescent peaks in the range



16.8 Room-temperature μ-PL spectrum from T6 microcavity. The pumping power is 0.2 mW at 514.5 nm. The reference spectrum at the bottom is recorded in the PhC pattern-free region.

of 1.3 to 1.6 μm dominate the spectrum and each peak represents one cavity mode supported by the PhC microcavity. The reference spectrum at the bottom of the figure is recorded in the PhC pattern-free region of the same sample under the same condition. Significant enhancement is achieved due to the optical resonance in the cavity. The Purcell effect may enhance the on-resonance luminescence greatly and suppress the spontaneous emission rate at the other wavelengths as shown in the figure. The full-width-half-maximum (FWHM) of the strongest peak located at 1.39 μm is 3.15 nm, which corresponds to a quality factor around 440, and the largest quality factor of 600 is observed for the peak located at 1.5 μm.

To study the dependence of the resonant peak wavelengths on the PhC lattice constant, T6 microcavities with different lattice constants were fabricated and tested by room-temperature PL measurements [22]. Figure 16.9 shows the room-temperature μ-PL spectra from the T6 microcavities with different PhC lattice constants. To clearly show the relationship between the luminescent peaks for different cavities, the scale of the wavelength axis of the spectra is decreased by the reciprocal ratio of the lattice constant, and the axis is shifted to match the peak positions. Comparing the five spectra, it is obvious that the resonant peaks move to the longer-wavelength side

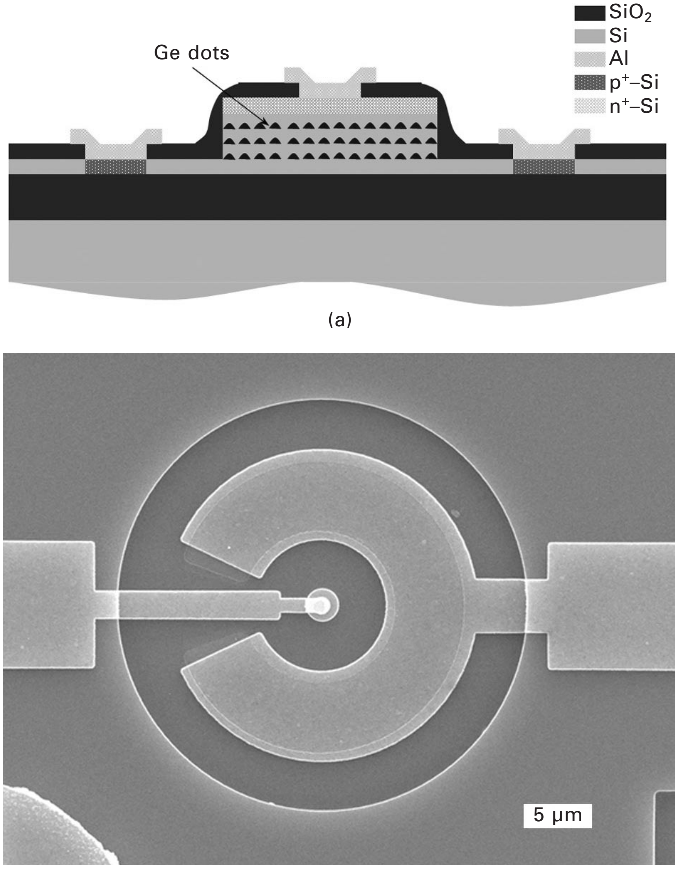


16.9 Room-temperature μ-PL spectra from T6 microcavities with different lattice constants. The scale of the wavelength axis is decreased and moved to clearly show the relationship between the spectra. The vertical dashed lines show the alignment of four selected resonant peaks.

when the lattice constant is increased. This is reasonable considering that the photonic bandgap of the PhC structure is proportional to the lattice constant of the PhC lattice. The vertical dashed lines indicate the peak positions of four selected peaks, and the peak positions coincide well among different cavities. This agreement comes from the scaling law of the PhC structure. Recently, a quality factor as high as 20,000 has already been achieved for the resonant light emission from Ge dots in photonic crystal line defect cavities [27].

16.2.4 Current-injected devices

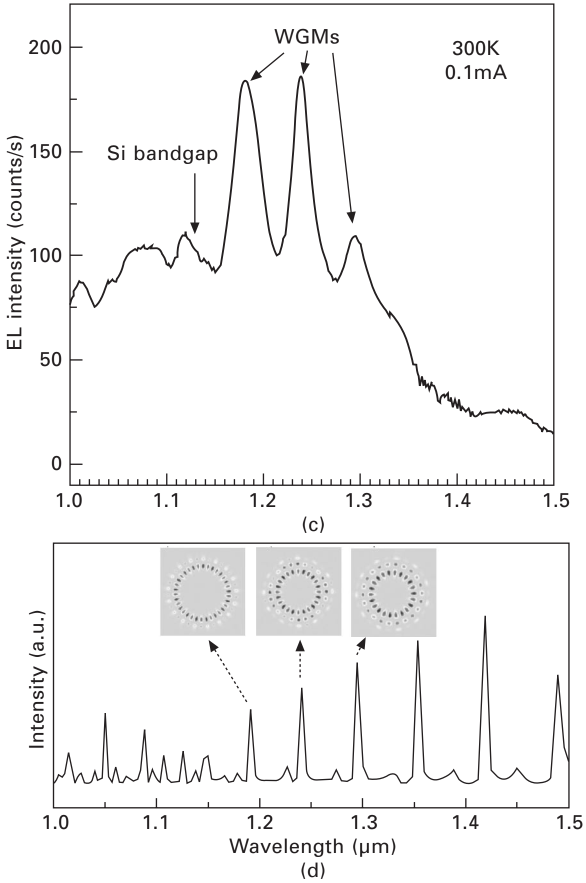
For practical devices, current injection is necessary in light-emitting devices (LEDs) based on Ge dots. Electroluminescence (EL) from SiGe heterostructures has been reported by several research groups [28, 29]. Room-temperature resonant EL from Si microdisks with Ge dots was reported recently [30]. Figures 16.10(a) and (b) show the schematic structure and SEM image of the fabricated device, respectively. Figure 16.10(c) shows the EL spectrum recorded at room temperature under an injected current of 0.1 mA. Clear EL is observed in the wavelength range of 1 to 1.4 μm. Several peaks are seen in the spectrum, and they may be assigned to the optical resonance in the microdisk. There are three major peaks, located at 1.185, 1.238, and 1.295 μm, respectively. Three-dimensional finite-difference-time-domain (FDTD)



(a)

(b)

16.10 (a) Schematic structure of the Ge dots microdisk EL device; (b) SEM image of a 2.8 μm microdisk EL device; (c) EL spectrum of the device under a current of 0.1 mA recorded at room-temperature; (d) calculated resonant wavelengths of TM-polarized-like WGMs supported by the microdisk. The insets show the mode profiles of the WGMs.



16.10 Continued

simulations identify these peaks. Figure 16.10(d) shows the calculated TM-polarized-like resonant peaks of the device. Three peaks located at 1.191, 1.241 and 1.296 μm correspond well to the three major peaks in the EL spectrum. The mode profiles of these peaks are shown by the insets in Fig. 16.10(d). They are whispering-gallery modes (WGMs) of order $\mathrm{TM0}_{2,\,27}$, $\mathrm{TM0}_{3,\,12}$, and $\mathrm{TM0}_{3,\,12}$. The demonstration of current-injected Si-based light-emitting devices based on Ge dots in optical microcavities shows a possible direction for the ultimate Si light source for hybrid integration on Si.

16.3 Silicon–germanium (SiGe) quantum cascade laser (QCL) structures

III–V semiconductor lasers covering the wavelength range from ultraviolet to far infrared are the most popular lasers in the market. However, efforts for a Si-based laser have been increasing in recent years due to its big potential for cost-effective monolithic optoelectronic integration on the Si platform. Quantum cascade laser (QCL) structures can solve the problem of the indirect bandgap of Si in principle, since the light emission is relying on the intersubband transitions. The energy of intersubband transitions in Si/SiGe cascade structures is small and independent of the bandgap of SiGe, resulting in light emission with a frequency in the terahertz (THz) range covering the mid- and far-infrared part of the electromagnetic spectrum. Several potential applications of THz lasers, including biomedical imaging [31, 32], gas sensing [33, 34], chemical spectroscopy, free-space communication [35, 36], and security monitoring, have driven QCLs into one of the major focuses of research. QCLs based on III–V semiconductors were realized in 1994 by researchers at Bell Laboratories [37], and much progress has been achieved in recent years for III–V semiconductor QCLs [38–41]. Since Si is the dominating material in the integrated circuit industry, pursuing Si/SiGe QCLs for optoelectronic integration on Si is natural and attractive due to the low cost, the safety, and the large production capacity of Si technology. The lack of strong polar optical phonon scattering in group IV materials is another advantage of SiGe QCL, which reduces the effect of temperature on the lifetime of carriers. No clear decrease of carrier lifetime due to polar optical phonon scattering has been observed in SiGe QC structures when the temperature is increased up to 100 K [42, 43].

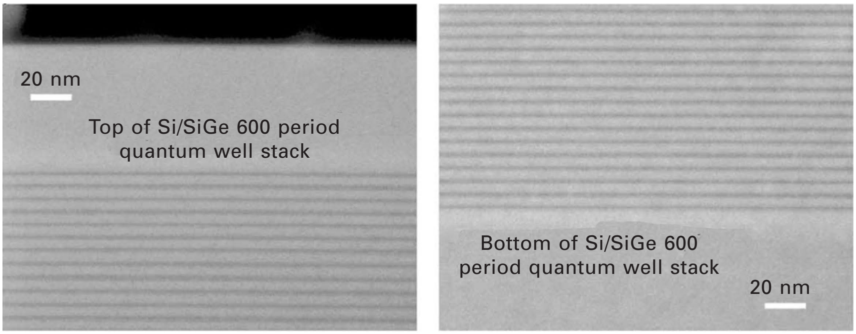
16.3.1 Growth of SiGe quantum cascade (QC) structures

The realization of a SiGe QCL is still difficult due to several obstacles including the strained material system, complicated band structures, and difficulties in doping. Since the conduction band offset of the Si/SiGe quantum well is small compared with that of a valence band, it is difficult to obtain significant quantum confinement for electrons in the quantum well. Moreover, the effective mass for electrons tunneling along the Si(100) growth direction is as large as $0.94m_e$, where m_e is the free electron mass [21, 44]. This means that very thin sub-nanometer barriers are necessary for efficient tunneling of electrons. But growth of such thin n-type layers is difficult since most of the n-type dopants show surface segregation. Therefore, most of the SiGe QC structures are designed to utilize intersubband transitions of holes in the valence band.

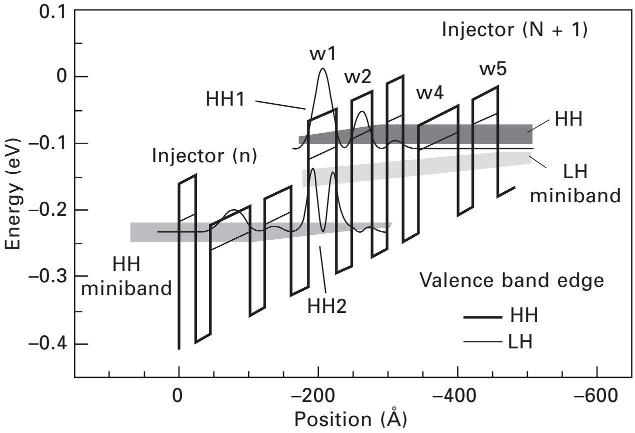
The difficulty induced by the 4.2% of lattice mismatch between Si and Ge is the main factor limiting the freedom of design of SiGe QC structures. Typical III–V QCLs have several hundred layers of thin films, but growth of a similarly complicated structure is difficult for SiGe systems. Usually, the critical thickness for strain relaxation is much smaller than the total growth thickness for QCL. Pseudomorphic growth of SiGe on Si wafer allows only a few periods with the thickness up to a metastable critical thickness. Strain symmetrization, which is achievable by careful design of the QC layering structure, allows the growth of a thick structure [44]. Each Ge-rich epilayer under compressive strain should be designed to be neighbored by a Si-rich epilayer under tensile strain. Therefore, the overall strain along the entire structure could be zero. Growth of 600 periods (over 4 μm thick) of a SiGe QC structure with perfect balanced strain has been demonstrated by Lynch and his colleagues [7] as shown in Fig. 16.11. There are no dislocations in the structure after 600 periods of growth and the surface is smooth without fluctuation. The ability to grow thick SiGe QC structures is essential for the design of SiGe QCLs.

16.3.2 Electroluminescence from SiGe QC structures

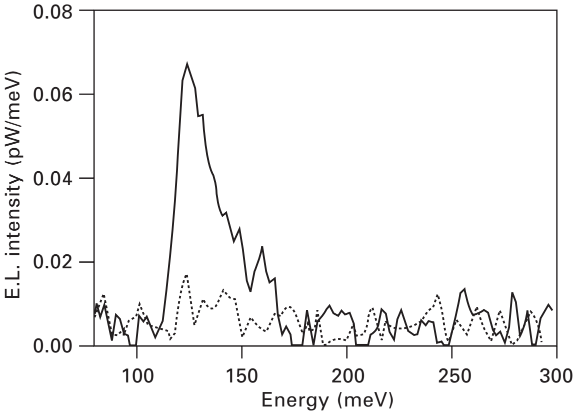
Intersubband EL from a Si/SiGe cascade structure at mid-infrared frequency was first demonstrated by Dehlinger and his colleagues in 2000 [45]. The QC structure was pseudomorphically grown on Si (0 0 1) substrate consisting of 12 highly strained cascades. The valence band potential structure at a bias of 50 kV/cm is shown in Fig. 16.12. The intersubband transitions between two heavy hole levels, HH1 and HH2, in the main well (W1) are the intended light emission. The carriers are injected into the upper level in the main well from the injector, and are confined without tuning out by the adjacent wells of W4 and W5. Figure 16.13 shows the intersubband light emission from the Si/SiGe QC structure, with the peak located at ~125 meV and the width of around 20 meV corresponding to the transition. As shown in Fig. 16.12, there is a light hole (LH) miniband between HH1 and HH2, which will induce scattering of holes from the upper level of HH2. The nonradiative HH–LH scattering is predicted to be very fast [46, 47], which reduces the lifetime of carriers on the HH2 level, resulting in great difficulty in achieving population inversion.


20 nm
Top of Si/SiGe 600 period quantum well stack
Bottom of Si/SiGe 600 period quantum well stack
20 nm

16.11 TEM figure of the top and bottom part of the cross-sectional view of a thick Si/SiGe QC structure. Adopted from Ref. 7.


Energy (eV)
0.1
0
–0.1
–0.2
–0.3
–0.4
0
–200
–400
–600
Position (Å)
Injector (N + 1)
w1
w2
w4
w5
HH1
HH
Injector (n)
LH
miniband
HH
miniband
HH2
Valence band edge
HH
LH

16.12 The potential structure of the valence band of a Si/SiGe QC structure at a bias of 50 kV/cm. Adopted from Ref. 45.



16.13 TM (solid line) and TE (dashed line) polarized EL from the Si/SiGe QC structure at 50 K. Adopted from Ref. 45.

In order to increase the lifetime of the upper HH2 level, Bormann and his colleagues use a modified structure to confine the HH2 and HH1 in separated wells to tune the overlap between them [48]. It is found that the lifetime of carriers on the HH2 level is increased when the width of the barrier between the two wells is larger than 2.5 nm. However, the weak coupling between HH2 and HH1 leads to a reduction in difficulty in lasing probability since the radiative transition probability between them is also decreased. The optimal width of the barrier of this kind of band design still remains to be addressed.

The propagation direction of electromagnetic emission originating from intersubband transitions between electron levels and HH levels in

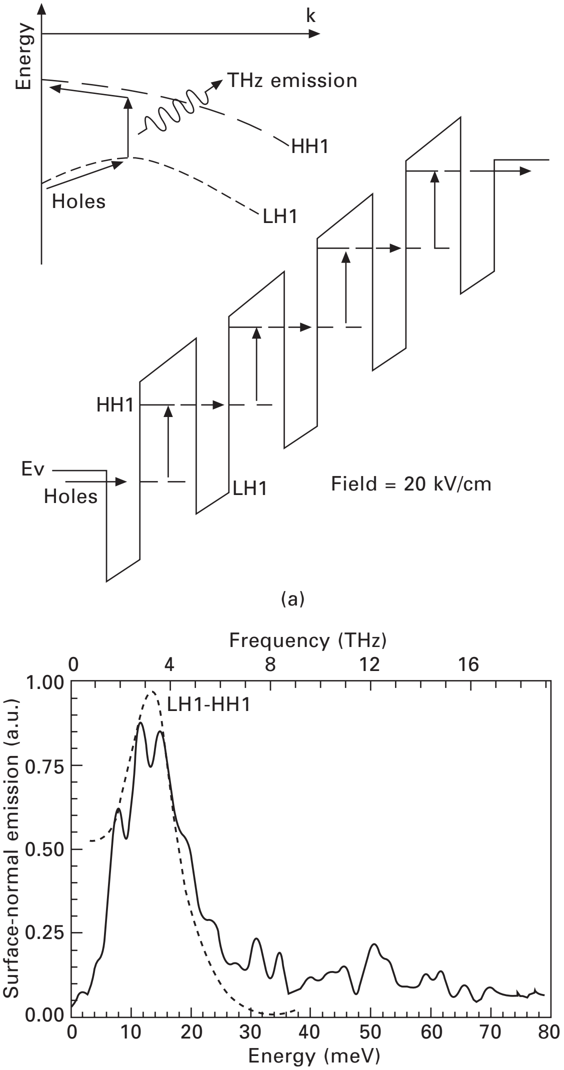
QC structures is parallel to the quantum well due to the quantization requirements. Therefore, it is difficult to fabricate vertical emitting devices based on HH to HH transitions. The situation is different for LH to HH intersubband transitions, since the momentum is changed in the transitions occurring between different p-states of the hole bands [49, 50]. Therefore, it is possible to fabricate vertical emitting devices utilizing the momentum breaking in the LH to HH intersubband transitions. SiGe QC emitters based on LH to HH intersubband transitions have been demonstrated by Paul and his colleagues [51, 52]. Figure 16.14(a) shows the schematic diagram of the quantum staircase emitter [51]. After the holes are injected into the point of $k = 0$, they are forced to finite k by the negative effective mass in the LH1 band, then the LH to HH intersubband transitions occur and then move back to the point of $k = 0$. Figure 16.14(b) shows the recorded EL from the QC emitter, and the LH–HH transition is clearly seen. The intersubband lifetime was measured to be around 30 ps and is relatively independent of the temperature, since it results from alloy scattering in the quantum well. This gives a significant advantage compared to III–V compound semiconductors in which the polar optical phonon scattering dominates the nonradiative transitions at temperatures higher than 40 K. The maximum output power is 10 nW at 60 K for a voltage of 2 V corresponding to an efficiency of 2.5×10^{-5}.

16.3.3 N-type SiGe QC structures

Recently, n-type SiGe/Si QCL has attracted some attention due to its simpler band structure and few types of intersubband transitions [53–56]. As shown in Fig. 16.15, Han and his colleagues proposed an n-type QC structure utilizing a Ge–SiGe superlattice, in which quantum wells for electrons are formed in the L and Γ valleys in the Ge layers [53, 54]. To make the electronic quantum well deep in the L valley, a $Ge/Si_{0.22}Ge_{0.78}$ superlattice was used as the active material. A vertical bound-continuum transition structure can be formed in the L valley under a bias of 80 kW/cm. The carriers in the lower level are extracted by miniband transport and L–Γ tunneling to the subband in the Γ well of the next period, and then the carriers are injected into the upper level in the L well by intervalley scattering. Numerical calculations show that population inversion is possible in the vertical structure at room temperature.

16.3.4 Optical waveguides for SiGe QCLs

One key issue for fabricating Si/SiGe QCLs is the design and fabrication of cavities to confine the THz light emission in optical waveguides. Unfortunately, the required minimum thickness for waveguides of THz light is about 15 μm

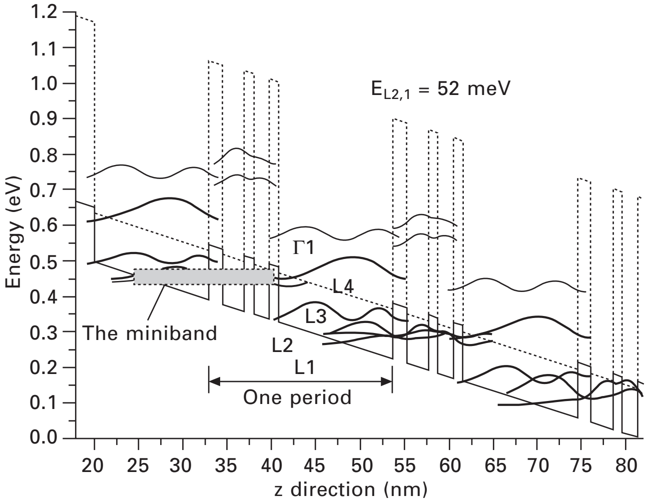


(a)

(b)

16.14 (a) The schematic hold band structure and diagram of the quantum staircase emitter based on LH to HH transitions. (b) Measured surface-normal EL at 60 K for the QC emitter. The dashed line is the calculated spectrum. Adopted from Ref. 51.

which is very difficult to achieve in epitaxy growth of the lattice-mismatched SiGe material. In order to confine the light in a thinner waveguide, surface plasmon modes have been successfully used in AlGaAs QCLs [57]. The top metal contacts provide a natural metal–semiconductor interface to pin the surface plasmon mode at the interface. At the bottom of the waveguide,



16.15 Conduction band energy diagram of an n-type Ge/SiGe QC structure under a bias of 80 kV/cm.

a highly doped semiconductor layer can provide partial confinement at the interface, leading to mode overlap of around 50% in AlGaAs QCLs and only 20% in SiGe QCLs [58]. In order to enhance the optical confinement, Kelsall and his colleagues used tungsten silicide (WSi_2) at the waveguide bottom to get over 95% mode confinement. The silicon-on-silicide substrates were prepared by bonding and an etch-back technique. EL at THz was observed at low temperature corresponding to interwell HH–LH transitions.

16.4 Conclusions

Silicon-based light sources in telecommunication wavelengths will be among the most important components for future optoelectronic hybrid integration on Si, which will make the next-generation hybrid VLSI a possible answer to keep Moore's law going for the next tens of years. Strong room-temperature resonant light emission from Ge self-assembled quantum dots in optical microcavities shows a promising way to fabricate fully CMOS-compatible silicon-based light emitting devices. It is highly attractive to the wide applied physics community, the photonics community, and also the IC industry. It will contribute to the ultimate silicon-based light source, optical interconnections inside Si chips, accelerating the emerging of the 'optical age of silicon'.

For the various applications based on THz and mid-infrared radiation, efficient and low-cost Si/SiGe QCLs are desired. Different Si/SiGe QC structures have been reported, and EL originating from both HH–HH and LH–HH intersubband transitions has been reported. Buried silicide technology may provide an answer to the requirement of vertical confinement of the THz emission in Si. However, and on the way to the ultimate Si/SiGe lasers, challenges still remain concerning various aspects, including structure design, epitaxy growth, and waveguide design. Along with continuous progress in these issues, Si/SiGe QCLs are expected to be realized in the near future and to boost the industries of biomedical imaging, gas sensing, chemical spectroscopy, and free-space communication.

16.5 References

1. D. K. Nayak, N. Usami, S. Fukatsu, and Y. Shiraki, 'Band-edge photoluminescence of SiGe/strained-Si/SiGe type-II quantum wells on Si(100)', *Appl. Phys. Lett.* 63(25), 3509 (1993)
2. S. Fukatsu, N. Usami, Y. Kato, H. Sunamura, Y. Shiraki, H. Oku, T. Ohnishi, Y. Ohmori, and K. Okumura, 'Gas-source molecular beam epitaxy and luminescence characterization of strained $Si_{1-x}Ge_x$/Si quantum wells', *J. Crystal Growth* 136(1–4), 315–321 (1994)
3. S. Fukatsu, H. Sunamuru, Y. Shiraki, and S. Komiyama, 'Phononless radiative recombination of indirect excitons in a Si/Ge type-II quantum dot', *Appl. Phys. Lett.* 71(2), 258 (1997)
4. K. Kawaguchi, M. Morooka, K. Konishi, S. Koh, and Y. Shiraki, 'Optical properties of strain-balanced SiGe planar microcavities with Ge dots on Si substrates', *Appl. Phys. Lett.* 81(5), 817 (2002)
5. M. W. Dashiell, U. Denker, C. Müller, G. Costantini, C. Manzano, K. Kern, and O. G. Schmidt, 'Photoluminescence of ultrasmall Ge quantum dots grown by molecular-beam epitaxy at low temperatures', *Appl. Phys. Lett.* 80(7), 1279 (2002)
6. E. S. Kim, N. Usami, and Y. Shiraki, 'Control of Ge dots in dimension and position by selective epitaxial growth and their optical properties', *Appl. Phys. Lett.* 72(13), 1617 (1998)
7. S. A. Lynch, D. J. Paul, P. Townsend, G. M. Matmon, Z. Suet, R. W. Kelsall, Z. Ikonic, P. Harrison, J. Zhang, D. J. Noriss, A. G. Cullis, C. R. Pidgeon, P. Murzyn, B. Murdin, M. Bain, H. S. Gamble, M. Zhao, and W. X. Ni, 'Toward silicon-based lasers for terahertz sources', *IEEE J. Selected Topics in Quantum Electronics* 2(6), 1570–1577 (2006)
8. G. Dehlinger, L. Diehl, U. Gennser, H. Sigg, J. Faist, K. Ensslin, D. Grützmacher, and E. Müller, 'Intersubband electroluminescence from silicon-based quantum cascade structures', *Science* 290, 2277–2280 (2000)
9. G. Franzò, F. Priolo, S. Coffa, A. Polman, and A. Camera, 'Room-temperature electroluminescence from Er-doped crystalline Si', *Appl. Phys. Lett.* 64(17), 2235 (1994)
10. L. Pavesi, L. Dal Negro, C. Mazzoleni, G. Franzò, and F. Priolo, 'Optical gain in silicon nanocrystals', *Nature* 408, 440–444 (2000)
11. W. L. Ng, M. A. Lourenço, R. M. Gwilliam, S. Ledain, G. Shao, and K. P. Homewood, 'An efficient room-temperature silicon-based light-emitting diode', *Nature* 410, 192–194 (2001)
12. E. M. Purcell, 'Spontaneous emission probabilities at radio frequencies', *Phys. Rev.* 69, 681 (1949)

13. E. Yablonovitch, 'Inhibited spontaneous emission in solid-state physics and electronics', *Phys. Rev. Lett.* 58, 2059–2063 (1987).
14. K. Kawaguchi, S. Koh, Y. Shiraki, and J. Zhang, 'Fabrication of strain-balanced $Si_{0.73}Ge_{0.27}$/Si distributed Bragg reflectors on Si substrates', *Appl. Phys. Lett.* 79(4), 476–478 (2001)
15. K. Kawaguchi, M. Morooka, K. Konishi, S. Koh, and Y. Shiraki, 'Optical properties of strain-balanced SiGe planar microcavities with Ge dots on Si substrates', *Appl. Phys. Lett.* 81(5), 817–819 (2002)
16. K. Kawaguchi, S. Koh, Y. Shiraki, and J. Zhang, 'Fabrication of strain-balanced $Si_{0.73}Ge_{0.27}$/Si-distributed Bragg reflectors on Si substrates for optical device applications', *Physica E* 13(2–4), 1051–1054 (2002)
17. K. Kawaguchi, K. Konishi, S. Koh, Y. Shiraki, Y. Kaneko, and J. Zhang, 'Optical properties of strain-balanced $Si_{0.73}Ge_{0.27}$ planar microcavities on Si substrates', *Jpn. J. Appl. Phys.* 41, 2664–2667 (2002)
18. S. L. McCall, A. F. J. Levi, R. E. Slusher, S.J. Pearton, and R.A. Logan, 'Whispering-gallery mode microdisk lasers', *Appl. Phys. Lett.* 60(3), 289 (1992)
19. J. S. Xia, K. Nemoto, Y. Ikegami, N. Usami, and Y. Shiraki, 'Silicon-based light emitters fabricated by embedding Ge quantum dots in Si microdisks', *Appl. Phys. Lett.* 91, 011104 (2007)
20. J. S. Xia, R. Tominaga, N. Usami, S. Iwamoto, Y. Ikegami, K. Nemoto, Y. Arakawa, and Y. Shiraki, 'Resonant photoluminescence from Ge self-assembled dots in optical microcavities', *J. Crystal Growth* 311, 883 (2009)
21. J. S. Xia, Y. Ikegami, Y. Shiraki, N. Usami and Y. Nakata, 'Strong resonant luminescence from Ge quantum dots in photonic crystal microcavity at room temperature', *Appl. Phys. Lett.* 89(20), 201102 (2006)
22. J. Xia, R. Tominaga, S. Fukamizu, N. Usami, and Y. Shiraki, 'Generation and wavelength control of resonant luminescence from silicon photonic crystal microcavities with Ge dots', *Jpn. J. Appl. Phy.* 48, 022102 (2009)
23. M. El Kurdi, S. David, P. Boucaud, C. Kammerer, X. Li, V. Le Thanh, S. Sauvage, and J.-M. Lourtioz, 'Strong 1.3–1.5 μm luminescence from Ge/Si self-assembled islands in highly confining microcavities on silicon on insulator', *J. Appl. Phys.* 96, 997–1000 (2004)
24. P. Boucaud, X. Li, M. El Kurdi, S. David, X. Checoury, S. Sauvage, C. Kammerer, S. Cabaret, V. Le Thanh, D. Bouchier, J.-M. Lourtioz, O. Kermarrec, Y. Campidelli, and D. Bensahel, 'Ge islands and photonic crystals for Si-based photonics', *Opt. Mater.* 27(5), 792–798 (2005)
25. X. Li, P. Boucaud, X. Checoury, O. Kermarrec, Y. Campidelli, and D. Bensahel, 'Probing photonic crystals on silicon-on-insulator with Ge/Si self-assembled islands as an internal source', *J. Appl. Phys.* 99, 023103 (2006)
26. X. Li, P. Boucaud, X. Checoury, M. El Kurdi, S. David, S. Sauvage, N. Yam, F. Fossard, D. Bouchier, J. M. Fédéli, A. Salomon, V. Calvo, and E. Hadji, 'Quality factor control of Si-based two-dimensional photonic crystals with a Bragg mirror', *Appl. Phys. Lett.* 88, 091122 (2006)
27. M. El Kurdi, X. Checoury, S. David, T. P. Ngo, N. Zerounian, P. Boucaud, O. Kermarrec, Y. Campidelli, and D. Bensahel, 'Quality factor of Si-based photonic crystal L3 nanocavities probed with an internal source', *Opt. Express* 16, 8780–8791 (2008).
28. S. Fukatsu, N. Usami, T. Chinzei, Y. Shiraki, A. Nishida, and K. Nakagawa, 'Electroluminescence from strained SiGe/Si quantum well structures grown by solid

source SI molecular beam epitaxy', *Jpn. J. Appl. Phys.* 31, L1015–L1017 (1992).

29. T. Brunhes, P. Boucaud, S. Sauvage, F. Aniel, J.-M. Lourtioz, C. Hernandez, Y. Campidelli, O. Kermarrec, D. Bensahel, G. Faini, and I. Sagnes, 'Electroluminescence of Ge/Si self-assembled quantum dots grown by chemical vapor deposition', *Appl. Phys. Lett.* 77, 1822–1824 (2000)
30. J. Xia, Y. Takeda, N. Usami, T. Maruizumi, and Y. Shiraki, 'Room-temperature electroluminescence from Si microdisks with Ge quantum dots', *Opt. Express* 18(13), 13945–13950 (2010)
31. J. Darmo, V. Tamosiunas, G. Fasching, J. Kröll, K. Unterrainer, M. Beck, M. Giovannini, J. Faist, C. Kremser, and P. Debbage, 'Imaging with a terahertz quantum cascade laser', *Opt. Express.* 12(9), 1879–1884 (2004)
32. D. D. Arnone, C. M. Ciesla, and M. Pepper, 'Terahertz imaging comes into view', *Phys. World* 13(4), 35–40 (2000)
33. C. Gmachl, F. Capasso, R. Kohler, A. Trediccui, L. Hutchinson, D. L. Sivco, J. N. Baillargeon, and A. Y. Cho, 'The sense-ability of semiconductor lasers', *IEEE Circuits and Devices*, 20, 10–17 (2000).
34. C. Roller, A. A. Kosterev, F. K. Tittel, C. Gmachl, and D. L. Sivco, 'Carbonyl sulfide detection with a thermoelectrically cooled midinfrared quantum cascade laser', *Opt. Lett.*, 28, 2052–2054 (2003)
35. J. Oulette, 'Quantum cascade lasers turn commercial', *Industrial Physicist* 8–13 (2001).
36. F. Capasso, R. Paiella, R. Martini, R. Colombelli, C. Gmachl, T. L. Myers, M. Taubm, R. Williams, C. Bethea, K. Unterrainer, H. Hwang, D. L. Sivco, A. Y. Cho, A. M. Sergent, H. C. Liu, and E. D. Whittaker, 'Quantum cascade lasers: ultrahigh-speed operation, optical wireless communication, narrow linewidth, and far infrared emission', *IEEE J. Select. Topics Quantum Electronics* 5, 511–529 (2002)
37. J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, 'Quantum cascade laser', *Science* 264, 553–556 (1994)
38. C. Sirtori, J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho, 'Quantum cascade laser with plasmon-enhanced waveguide operating at 8.4 μm wavelength', *Appl. Phys. Lett.* 66(24), 3242 (1995)
39. R. Colombelli, F. Capasso, C. Gmachl, A. L. Hutchinson, D. L. Sivco, A. Tredicucci, M. C. Wanke, A. M. Sergent, and A. Y. Cho, 'Far-infrared surface-plasmon quantum-cascade lasers at 21.5 μm and 24 μm wavelengths', *Appl. Phys. Lett.* 78(18), 2620 (2001)
40. M. Beck, D. Hofstetter, T. Aellen, J. Faist, U. Oesterle, M. Ilegems, E. Gini, and H. Melchior, 'Continuous wave operation of a mid-infrared semiconductor laser at room temperature', *Science* 295, 301–305 (2002)
41. S. Kumar, B. S. Williams, S. Kohen, Q. Hu, and J. L. Reno, 'Continuous-wave operation of terahertz quantum-cascade lasers above liquid-nitrogen temperature', *Appl. Phys. Lett.* 84(14), 2494 (2004)
42. C. R. Pidgeon, P. Murzyn, J.-P. R. Wells, I. V. Bradley, Z. Ikonic, R. W. Kelsall, P. Harrison, S. A. Lynch, D. J. Paul, D. D. Arnone, D. J. Robbins, D. Norris, and A. G. Cullis, 'THz intersubband dynamics in p-Si/SiGe quantum well structures', *Physica E* 13(2–4), 904–907 (2002)
43. P. Murzyn, C. R. Pidgeon, J.-P. R. Wells, I. V. Bradley, Z. Ikonic, R. W. Kelsall, P. Harrison, S. A. Lynch, D. J. Paul, D. D. Arnone, D. J. Robbins, D. Norris, and A. G. Cullis, 'Picosecond intersubband dynamics in p-Si/SiGe quantum-well emitter structures', *Appl. Phys. Lett.* 80(8), 1456 (2002)

44. D. J. Paul, 'Si/SiGe heterostructures: From materials and physics to devices and circuits', *Semicon. Sci. Technol.* 19(10), R75–108 (2004)
45. G. Dehlinger, L. Diehl, U. Gennser, H. Sigg, J. Faist, K. Ensslin, D. Grützmacher, and E. Müller, 'Intersubband electroluminescence from silicon-based quantum cascade structures', *Science* 290, 2277–2280 (2000)
46. T. Elsaesser and M. Woerner, 'Femtosecond infrared spectroscopy of semiconductors and semiconductor nanostructures', *Phys. Rep.* 321(6), 253–305 (1999)
47. R. A. Kaindl, M. Wurm, K. Reimann, M. Woerner, T. Elsaesser, C. Miesner, K. Brunner, and G. Abstreiter, 'Ultrafast dynamics of intersubband excitations in a quasi-two-dimensional hole gas', *Phys. Rev. Lett.* 86, 1122–1125 (2001)
48. I. Bormann, K. Brunner, S. Hackenbuchner, G. Abstreiter, S. Schmult, and W. Wegscheider, 'Nonradiative relaxation times in diagonal transition Si/SiGe quantum cascade structures', *Appl. Phys. Lett.* 83(26), 5371 (2003)
49. L. Friedman, G. Sun, and R. A. Soref, 'SiGe/Si THz laser based on transitions between inverted mass light-hole and heavy-hole subbands', *Appl. Phys. Lett.* 78(4), 401 (2001)
50. G. Sun, Y. Lu, and J. B. Khurgin, 'Valence intersubband lasers with inverted light-hole effective mass', *Appl. Phys. Lett.* 72(12), 1481 (1998)
51. D. J. Paul, S. A. Lynch, R. Bates, Z. Ikonic, R. W. Kelsall, P. Harrison, D. J. Norris, S. L. Liew, A. G. Cullis, D. D. Arnone, C. R. Pidgeon, P. Murzyn, J.-P. R. Wells, and I. V. Bradley, 'Si/SiGe quantum-cascade emitters for terahertz applications', *Physica E* 16, 147–155 (2003)
52. R. Bates, S. A. Lynch, D. J. Paul, Z. Ikonic, R. W. Kelsall, P. Harrison, S. L. Liew, D. J. Norris, A. G. Cullis, W. R. Tribe, and D. D. Arnone, 'Interwell intersubband electroluminescence from Si/SiGe quantum cascade emitters', *Appl. Phys. Lett.* 83(20), 4092 (2003)
53. G. Han, J. Yu, and Y. Liu, 'n-Type Ge–SiGe quantum cascade structure utilizing quantum wells for electrons in the L and Γ valleys', *IEEE Photon. Technol. Lett.* 20(6), 419 (2008)
54. G. Q. Han and J. Z. Yu, 'An n-type SiGe/Ge QC structure utilizing the deep Ge quantum well for electron at the G point', *Semicond. Sci. Technol.* 22(7), 769–773 (2007)
55. K. Driscoll and R. Paiella, 'Silicon-based injection lasers using electronic intersubband transitions in the L valleys', *Appl. Phys. Lett.* 89(19), 191110 (2006)
56. Z. Ikonic, I. Lazic, V. Milanovic, R. W. Kelsall, D. Indjin, and P. Harrison, 'N-Si/SiGe quantum cascade structures for THz emission', *J. Lumin.* 121(2), 311–314 (2006)
57. R. Köhler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, R. C. Iotti, and F. Rossi, 'Terahertz semiconductor-heterostructure laser', *Nature* 417, 156–159 (2002)
58. Z. Ikonić, R. W. Kelsall, and P. Harrison, 'Waveguide design for mid- and far-infrared p-Si/SiGe quantum cascade lasers', *Semicond. Sci. Technol.* 19(1), 76 (2004)

17

Silicide and germanide technology for interconnections in ultra-large-scale integrated (ULSI) applications

S. ZAIMA and O. NAKATSUKA, Nagoya University, Japan

Abstract: This chapter discusses the silicide and germanide technology for interconnects in ULSI applications. After a brief introduction about silicide and contacts for ULSI, the formation of silicides and germanides on Si and SiGe is reviewed. Then, the chapter discusses crystalline properties of silicides and the effect of incorporation of third materials, such as Ge, Pt, and Pd, into silicide. The electrical properties of metal/Si, SiGe and Ge contacts are also reviewed.

Key words: silicide, germanide, contact, Schottky barrier height (SBH), contact resistance.

17.1 Introduction

In this chapter, we will introduce silicide and germanide technology and discuss their effects on Si and SiGe for interconnects such as metal/semiconductor contacts and metal gates in ultra-large-scale integrated (ULSI) applications. We will also summarize their formation processes and their crystalline and electrical properties.

Silicide and germanide are intermetallic compounds with Si and Ge, respectively. Many metals form stable silicides. The interfacial reactions between various metals and Si or Ge have been investigated for many years and some silicides and germanides have been developed for contact and gate materials in Si ULSI applications [1–8].

The application of metal silicide for Si metal-oxide-semiconductor field-effect transistors (MOSFETs) provides the low series resistance of the poly-Si gate and the reduction of contact resistances in the source/drain (S/D) regions. The following features are required for silicides in ULSI applications: low resistivity for low sheet resistance, low contact resistivity with Si for the reduction of parasitic resistance, high thermal robustness of silicide films during the fabrication processes of ULSIs, and low reactivity to prevent undesired reactions with interconnect materials such as W, Cu, and Al, or Si. According to the technology roadmap [9], development of ultra-thin silicide layers below 20 nm on ultra-shallow S/D junctions and ultra-low contact resistivity as low as 10^{-8} Ω cm^2 are required.

17.2 Formation of silicide and germanosilicide thin films

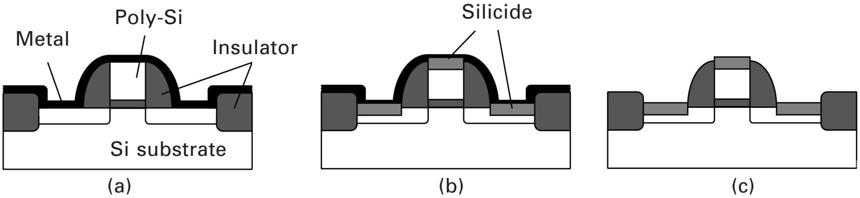
The self-aligned silicide (salicide) process has been one of the most important technologies for the shrinkage of MOSFETs. Figure 17.1 shows a schematic diagram of the salicide formation in MOSFETs.

A silicide film is formed by the deposition of a metal thin film on a Si substrate and the solid-phase reaction with post-deposition annealing (PDA). Silicide is only formed between metal and Si on poly-Si gates and on S/D contacts where metal comes directly into contact with Si. Metal deposited on insulators such as SiO_2 or Si_3N_4 either does not react or hardly reacts. Unreacted metals on these insulator areas can be selectively removed by wet chemical solutions after the silicidation, and salicide remains only at the gate and contact regions. This salicide process prevents the shortening of the gate to the S/D electrodes without the use of any lithography and patterning processes and minimizes the overlaps of these electrodes.

For CMOS applications, titanium disilicide ($TiSi_2$), cobalt disilicide ($CoSi_2$), and nickel monosilicide (NiSi) have been extensively investigated.

A layer of the C54-$TiSi_2$ phase is generally formed for contacts by the salicide process with two-step annealing. In the first step, the C49-$TiSi_2$ phase is formed by annealing in a N_2 atmosphere at a temperature range between 600°C and 700°C. This annealing suppresses the lateral diffusion of Si due to the TiN formed on the surface and prevents the bridging of silicide between the gate and S/D regions. After the selective etching of TiN and unreacted Ti, the C49 phase is transformed to the C54 phase by the second step of annealing at higher than 800°C.

$CoSi_2$ can also be formed by a salicide process with two-step annealing, much like C54-$TiSi_2$. The first step, annealing, provides a Co-rich-phase silicide such as Co_2Si or CoSi at a low temperature ranging between 400–500°C. After chemical selective etching of unreacted Co on insulators, $CoSi_2$, the



17.1 A schematic diagram of salicide formation in MOSFETs: (a) after the deposition of a metal layer, (b) after silicidation annealing, and (c) after selective etching of an unreacted metal layer on insulator.

low-resistivity phase, is formed by the second step of annealing at a high temperature, ranging between 600–800°C.

The formation mechanism of Co silicides is quite complicated, because diffusion species during Co silicide formation are dependent upon formed silicide phases. In the first stage of Co silicidation, the metal-rich Co_2Si layer grows predominantly by the diffusion of Co atoms during the Co/Si interfacial reaction, and then the CoSi layer grows predominantly by the diffusion of Si atoms into the Co_2Si layer [10–14]. In the Si-rich phase stage ($CoSi_2$ formation), the nucleation of $CoSi_2$ takes place at three points between CoSi grains and the Si substrate, and the $CoSi_2$ layer grows predominantly by the diffusion of Co atoms.

In general, the formation temperature of NiSi is generally lower than that of $TiSi_2$ and $CoSi_2$, and the trend with recent ULSI technology is that it is preferable to lower the process temperature. Additionally, the dominant diffusion species in NiSi formation is the metal Ni, while Si dominantly diffuses in the formation of Ti and Co silicides. This is an advantage for the formation of contacts in comparison with Ti and Co silicides, because the low diffusivity of Si into metal leads to the avoidance of silicide overgrowth on the insulator around the contact. The generation of vacancies in the Si diffusion layer is also suppressed during contact formation.

Recently, the formation of a Si–Ge–metal ternary compound (germanosilicide) on $Si_{1-x}Ge_x$ or Ge has become necessary, because of applications with $Si_{1-x}Ge_x$ S/D structures or high-mobility Ge channels.

In the case of the Ti/$Si_{1-x}Ge_x$/Si systems, a titanium digermanosilicide, $Ti(Si_{1-x}Ge_x)_2$, with the same ratio of Si and Ge as the $Si_{1-x}Ge_x$ layers, is initially formed [15, 16]. However, the additional annealing leads to the segregation of Ge from $Ti(Si_{1-x}Ge_x)_2$ and the formation of Ge-rich $Si_{1-z}Ge_z$ polycrystalline grains at germanosilicide grain boundaries [17, 18]. These phenomena can be explained by the lower heat for the formation of C54-$TiSi_2$ (-57.0 kJ mol^{-1}) than for C54-$TiGe_2$ (-47.5 kJ mol^{-1}) [18].

In the case of Ni/$Si_{1-x}Ge_x$/Si systems, the formation of Ni monogermanosilicide, $Ni(Si_{1-x}Ge_x)$, exhibits the same tendency [19, 20]. After annealing at a temperature as low as 350°C, the Ge fractions of $Ni(Si_{1-x}Ge_x)$ are close to those of the $Si_{1-x}Ge_x$ layers. The Ge fractions of $Ni(Si_{1-x}Ge_x)$ decrease to zero with an increase in the temperature of additional annealing. NiSi and NiGe phases have the same orthorhombic MnP-type crystal structure and the formation temperature of NiGe (260°C) is slightly lower than that of NiSi (350°C). On the other hand, the heat for the formation of NiSi (-42 kJ mol^{-1}) is negatively larger than that of NiGe (-32 kJ mol^{-1}), which means that NiSi is energetically more stable than NiGe [21]. Therefore, Ge atoms are included in the $Ni(Si_{1-x}Ge_x)$ phase formed at a lower temperature and they out-diffuse from germanosilicide to form the thermodynamically stable phase, NiSi.

17.3 Crystalline properties of silicides

In this section, the crystalline properties of major silicides for ULSI contacts, such as $TiSi_2$, $CoSi_2$, and NiSi, will be summarized. Table 17.1 summarizes some important crystalline and electrical properties of these silicides [2, 20].

17.3.1 $TiSi_2$

$TiSi_2$ is one of the most popular contact materials for ULSI applications in refractory metal silicides. It has high thermal stability and low resistivity. $TiSi_2$ has two types of crystalline structures: a base-centered orthorhombic C49 structure and a face-centered orthorhombic C54 structure. C49- and C54-$TiSi_2$ are high-resistivity (~60 μΩ cm) and low-resistivity phases (13–20 μΩ cm), respectively [21–24].

The sheet resistance of a $TiSi_2$ layer significantly increases when the width of a $TiSi_2$ line is shrunk. This is called the 'narrow line effect'. This narrow line effect restricts the application of $TiSi_2$ to nanoscale devices. The phase transformation from the C49 to the C54 phase does not take place for narrower poly-Si gate lines or smaller S/D contact areas below 500 nm and 100 μm^2, respectively [25, 26]. The reason is considered to be that the nucleation mode of the C54 phase is dependent upon the size of $TiSi_2$ layers [21, 27]. The energy barrier for the silicide nucleation increases by decreasing the layer thickness, because a disk-shaped C54 nucleus in a very thin layer has a larger surface-to-volume ratio than the spherical-shaped C54 nucleus in a thick layer. Hence, the nucleation of the C54 phase occurs preferentially at the three points of C49 grain boundaries for silicide layers thinner than 55 nm, while the nucleation of the C54 phase occurs predominantly at the grain boundary of the two C49 grains for layers thicker than 100 nm. As a result, the transformation of C54-$TiSi_2$ is limited depending on the size

Table 17.1 Summary of crystalline and electrical properties of C54-$TiSi_2$, $CoSi_2$, and NiSi

	C54-$TiSi_2$	$CoSi_2$	NiSi
Crystalline structure	Face-centered orthorhombic	Cubic	Orthorhombic
Formation temperature (°C)	700	700	350
Melting point (°C)	1500	1326	992
Si consumption per nm of metal (nm)	2.27	3.64	1.83
Thin film resistivity (Ω μm)	13–20	14–20	14–20
Dominant diffusion species for silicide formation	Si	Co and Si	Ni
SBH for n-type Si (eV)	0.49	0.61	0.68
SBH for p-type Si (eV)	0.51	0.43	0.38

of C49-$TiSi_2$ layers, and the resistance of $TiSi_2$ lines increases with the narrowing of the line width below 1 µm [21, 27, 28].

17.3.2 $CoSi_2$

$CoSi_2$ is an alternative material to C54-$TiSi_2$ for ULSI applications [29, 30] because $CoSi_2$ has low resistivity and high thermal robustness like C54-$TiSi_2$. Note that the original advantage of $CoSi_2$ was its insensitivity to decreasing line width compared to C54-$TiSi_2$. However, it has recently been reported that $CoSi_2$ thin layers also exhibit a scaling issue, namely, that the resistance of $CoSi_2$ lines increases with narrowing below 200 nm [31, 32]

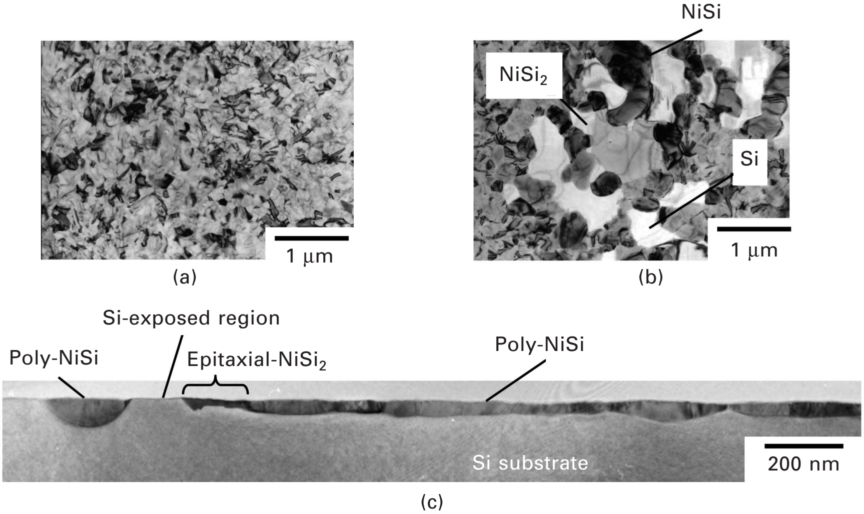
17.3.3 NiSi

NiSi is one of the promising candidates for silicide/Si contact materials alternative to Ti and Co silicides in next-generation ULSI devices [33–37]. NiSi/Si contacts have several advantages, as follows: (1) the low resistivity of NiSi, as low as that of C54-$TiSi_2$ and $CoSi_2$, (2) the low Si consumption during silicide formation compared to disilicides like $TiSi_2$ and $CoSi_2$, which is suitable for ultra-shallow junctions, (3) the low formation temperature of NiSi (below 400ºC), and (4) the low contact resistivity, as low as 10^{-8} Ω cm^2.

However, the poor thermal stability of NiSi layers is a critical issue for device applications, as it causes the degradation of the electrical properties such as increasing sheet resistance after the annealing processes [33–37]. In fact, the agglomeration of a NiSi thin layer easily occurs when annealing over 650°C, and the transformation of NiSi to the high resistivity phase, $NiSi_2$, takes place by annealing over 750°C. An epitaxial $NiSi_2$ layer is formed on Si(001) by the solid-phase reaction of a Ni/Si system. However, the interface morphology is generally not superior for contacts, due to the inhomogeneous formation of $NiSi_2$ and {111} facets at the epitaxial $NiSi_2$/Si interface, because the nucleation of epitaxial $NiSi_2$ occurs preferentially in the vicinity of the Si exposed region and agglomerated NiSi grains, as shown in Fig. 17.2 [33, 38].

The incorporation of a third element into a silicide/Si system is one of the effective solutions to improve the properties of silicide films, to control the solid-phase reaction, raise thermal stability, and lower the film and contact resistances. Several elements that have been incorporated into Ni/Si systems to improve the properties of NiSi/Si contacts have been reported, such as Pt [39–41], Pd [42], Ge [20, 43–45], C [20, 46], N [47, 48] and H [49]. Some examples are introduced below.

First, the impact of Ge on the stability of Ni silicides/Si contacts is described. The thermal phase stability of NiSi can be improved by the formation of



17.2 Plan-view transmission electron microscopy images of a NiSi layer on Si: (a) after silicidation at 350°C, and (b) after additional annealing at 750°C for 180 s. (c) Cross-sectional scanning electron microscopy image of a NiSi/Si contact after annealing at 750°C for 180 s.

germanosilicide on $Si_{1-x}Ge_x$ [20, 44, 45]. While the phase transformation from NiSi to $NiSi_2$ takes place with annealing at a temperature as high as 800°C, a monogermanosilicide phase formed on $Si_{1-x}Ge_x$ layers remains even after annealing at 850°C [20]. In a Ni–Ge system, the germanium-rich phase is not formed at any temperature, contrary to the Ni–Si system. Hence, when the Ni-monogermanosilicide phase transforms to the $NiSi_2$ phase in a Ni–Si–Ge system, an additional thermally-activated process is required, i.e., the segregation of Ge atoms from the germanosilicide [50]. This leads to the raising of the phase transformation temperature, in contrast to cases without Ge. However, the agglomeration of Ni germanosilicide is generally more severe compared to NiSi/Si systems [20].

The Pt incorporation into NiSi is effective for suppressing the agglomeration of a polycrystalline NiSi layer. Pt in NiSi affects the texture structure of a NiSi layer on a Si substrate [39, 51]. Crystallite grains in polycrystalline NiSi layers exhibit five different types of preferential orientation on a Si(001) substrate. These texture components change and the epitaxial alignment of NiSi on Si becomes preferable due to the expansion of unit cells caused by the addition of Pt [51]. The expansion of the NiSi lattice with Pt also causes the transformation of a MnP-type structure (orthorhombic) into a more stable NiAs-type structure (hexagonal). As a result, NiSi layers with Pt on a Si substrate become more stable above 800°C.

Moreover, the Pt incorporation into NiSi improves on the thermal stability of the low-resistivity NiSi phase. Since the crystalline structure of PtSi is a MnP-type like NiSi, and the lattice mismatch between NiSi and PtSi is less than 15%, NiSi and PtSi form a solid solution in bilayer and concentrated alloy thin layers. Pt–Si compounds do not form a disilicide phase, in contrast to a Ni–Si system, and the solid solubility of Pt in $NiSi_2$ may be low. Hence, the formation of the $NiSi_2$ phase can be suppressed by the Pt incorporation because of the required expulsion of Pt for $NiSi_2$ nucleation from NiSi.

The incorporation of Pt is also effective for improving the thermal stability of $Ni(Si_{1-x}Ge_x)$ on $Si_{1-x}Ge_x$. The agglomeration of $Ni(Si_{1-x}Ge_x)$ can be suppressed by the incorporation of as much as 5% of Pt and leads to the low sheet resistance of Ni germanosilicide even after annealing at 700ºC [52, 53].

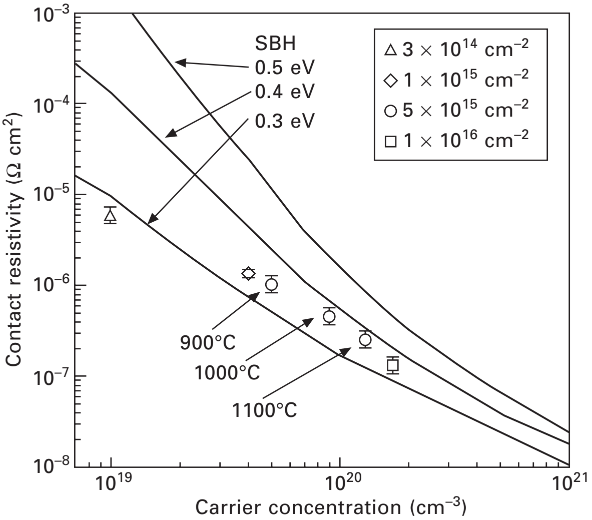
It is known that Pd also plays a role in raising the phase transition temperature from NiSi to $NiSi_2$ [42]. PdSi has the same crystalline structure (MnP-type) as that of NiSi and is generally formed above 700°C, a much higher formation temperature than that of NiSi. There is no Si-rich silicide in the Pd–Si system as well as the Pt–Si system. Therefore, the phase transformation temperature to $NiSi_2$ from NiSi is raised due to the introduction of PdSi into the NiSi layer.

The incorporation of C into Si substrates effectively improves the thermal stability and the electrical properties of NiSi/Si contacts [20, 46]. The increase in sheet resistance of NiSi layers on Si can be suppressed by the C implantation into Si substrates before silicidation, which is due to the prevention of the agglomeration of polycrystalline NiSi. The contact resistivity of NiSi/p^+-Si contacts with C is also reduced due to the pile-up of B atoms at the NiSi/Si interface, which is enhanced by the C incorporation.

17.4 Electrical properties

The contact resistivity of an ohmic metal/semiconductor contact is essentially dependent upon the Schottky barrier height (SBH) at the metal/Si interface and the impurity concentration in semiconductors. The theoretical estimation of contact resistivity has been reported [54, 55]. Figure 17.3 shows the contact resistivities of $CoSi_2$/Si contacts as a function of the impurity concentration. Contact resistivities numerically calculated for SBHs of 0.3, 0.4, and 0.5 eV are also shown in Fig. 17.3. As exhibited by this figure, a lower SBH and a higher impurity concentration are required to reduce contact resistivity. However, there are some limitation factors to achieving much lower contact resistivity, such as the energy bandgap, the electron affinity and the solid solubility of impurity atoms in Si.

In particular, it is difficult to simultaneously minimize the contact resistivities for both n^+- and p^+-Si by using the same contact material in

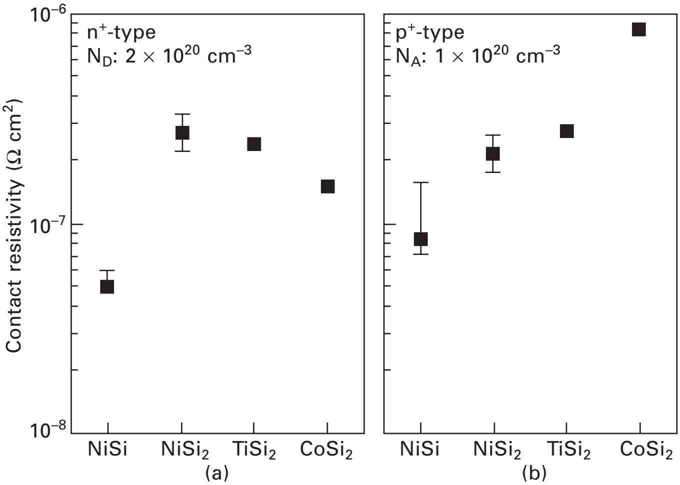


17.3 Numerically calculated contact resistivities of p-type contact as a function of the impurity concentration in Si for SBHs of 0.3, 0.4, and 0.5 eV. Experimental results of contact resistivity for $CoSi_2$/Si contacts are also shown.

CMOS devices, because the sum of SBHs for n$^+$- and p$^+$-Si is equal to the bandgap of Si; a material with a low contact resistivity for one type of Si is disadvantageous for another type. The incorporation of a SiGe layer into the silicide/Si interface is attractive to solve this dilemma, as will be described later.

Figures 17.4(a) and 17.4(b) show the contact resistivities of NiSi, $NiSi_2$, $CoSi_2$, and C54-$TiSi_2$/Si contacts for n$^+$- and p$^+$-types, respectively [20]. The contact resistivity of NiSi shows the lowest values for both n$^+$- and p$^+$-type contacts in these silicides. This is mainly attributed to the pile-up of the dopant impurity, P, at the NiSi/n$^+$-Si interface during salicide formation, and to the low SBH for the NiSi/p-type contact.

The junction leakage is also a serious issue when using $CoSi_2$ and NiSi for ohmic contacts on shallow S/D junctions of nanoscale MOSFETs. This problem becomes more serious in Co and Ni silicide contacts, because of the large diffusion coefficient of Co and Ni in Si compared to Ti [56]. The formation of $CoSi_x$ and $CoSi_2$ spikes in $CoSi_2$/Si contacts causes the junction leakage [57]. Additionally, Co atoms dissolve in Si substrates by post-annealing of very thin $CoSi_2$ islands on Si(001) above 600°C [58]. In-diffusion of Co and Ni atoms into a Si substrate during the formation of silicide/Si contacts, as well as the formation of generation–recombination centers, are suspected to be the causes of the intrinsic leakage current at shallow S/D regions [59, 60]. Pre-silicide implantation of As, Ge, F, and



17.4 Contact resistivities of NiSi, $NiSi_2$, $CoSi_2$, and C54-$TiSi_2$ contacts for (a) n^+- and (b) p^+- type Si. Impurity concentrations in n^+- and p^+-Si are 2×10^{20} cm^{-3} and 1×10^{20} cm^{-3}, respectively.

so on using an ion-implantation technique is effective for reducing the in-diffusion of Co and leads to the suppression of the S/D leakage current [61, 62].

The application of $Si_{1-x}Ge_x$ in a metal/Si ohmic contact has the following advantages: (1) controllability of the energy bandgap, (2) the possibility of realizing higher doping concentrations of impurities than those in Si, (3) the heteroepitaxial growth of $Si_{1-x}Ge_x$ on Si(001) substrates, and (4) good compatibility with Si processes. Since the energy bandgap of $Si_{1-x}Ge_x$ can be changed from 1.12 to 0.66 eV by increasing the Ge fraction, x, and the band offset is mainly formed at the valence band edge interfaces [63], we can expect smaller SBHs at the metal/p-$Si_{1-x}Ge_x$ interface than at the metal/Si interface. Also, smaller SBHs at the metal/semiconductor interface promise lower contact resistivity. A smaller SBH of the $TiSi_2$/SiGe/Si contact than the $TiSi_2$/Si contact has been reported [64]. In addition, the application of selectively epitaxial-grown $Si_{1-x}Ge_x$ and $Si_{1-x}C_x$ at S/D regions has recently become attractive for the realization of high-mobility channels due to the strain structure [65, 66].

Schottky barrier S/D is also an attractive contact structure for future MOSFETs. Schottky S/D provides an improvement on the device properties because it reduces parasitic capacitance and series resistance, minimizes the short-channel effect, is immune to the latch-up phenomenon, and suppresses the floating-body effect in SOI MOSFETs. This structure also leads to a simplification of the fabrication process due to its being free from the lithography limitation of S/D junction formation and the elimination of high-temperature annealing to form S/D junctions [67].

The lowering of SBHs is required to lower contact resistivity, and novel metal and rare earth silicides are some of the candidates for next-generation contact materials. Table 17.2 shows a summary of some of the properties of Pd, Pt, Ir, and Er silicides. $PdSi_2$, PtSi, and IrSi are expected for p-type contact materials because of the low SBH for p-type Si, and rare earth silicides such as $ErSi_2$ are expected for n-type contacts.

Recently, metal gates have been expected to replace poly-Si gates in MOSFETs in order to reduce the effective gate oxide thickness by eliminating the poly-Si depletion region. A fully silicided (FUSI) gate using, for example, Ni silicides is one promising candidate for metal gates [68, 69]. Silicides having the appropriate work functions for metal gates must be selected for the optimum operation of each n- and p-type MOSFET. The effective work function of Ni silicides strongly depends on silicide phases, the kinds of gate dielectrics and the dopant impurities in poly-Si before silicidation [68, 70, 71]. The Fermi level pinning (FLP) at the metal/dielectric interface is also a serious issue. It is reported that using a Ni-rich silicide like Ni_3Si on HfSiON provides unpinning of the Fermi level [71]. The formation of controlled silicide/dielectric interfaces in nanoscale gate regions with the appropriate work functions for CMOS applications remains a critical issue.

A Ge-channel MOSFET is an attractive alternative to a Si MOSFET for future ultra-large integrated circuits because both electron and hole mobility in Ge are higher than those in Si [72, 73]. It is necessary to establish metal/Ge contact technology for Ge MOSFET in order to realize high-performance Ge devices. SBH engineering at a metal/Ge interface is a key factor for realizing a low contact resistance for Ge MOSFET. The SBH for an n-type contact, ϕ_{Bn}, is generally formulated as

$$\phi_{Bn} = S(\phi_m - \chi_s) + \text{const.} \qquad 17.1$$

Table 17.2 Summary of crystalline and electrical properties of Pd, Pt, Ir, and Er silicides

	Pd_2Si	PtSi	IrSi	$ErSi_2$
Formation temperature (°C)	800	800	400	350
Melting point (°C)	1331	1229	–	1873
Si consumption per nm of metal for silicide formation (nm)	0.68	1.32	1.99	–
Resulting silicide thickness per nm of metal (nm)	1.69	1.98	3.57	–
Thin film resistivity (Ω µm)	14–35	28–35	–	30
Dominant diffusion species for silicide formation	Pd, Si	Pt	Si	Si
SBH for n-type Si (eV)	0.75	0.88	0.93	0.29
SBH for p-type Si (eV)	0.35	0.21	0.18	0.74

where ϕ_m is the work function of a metal and χ_s is the electron affinity of a semiconductor [67]. The index of interface behavior, *S*, is

$$S = \partial\phi_{Bn}/\partial\phi_m \tag{17.2}$$

It is known that the *S* value for metal/Ge contacts is very low ($S \sim 0.03$) compared to that for Si contacts ($S = 0.27$) and SBHs in various metal/n-type Ge contacts show values higher than 0.5 eV because of FLP at the metal/Ge interface [74, 75]. Many reports have discussed the origins of FLP, such as interface states due to metal-induced gap states (MIGS) or disorder-induced gap states (DIGS); however, the details regarding FLP in metal/Ge contacts have not yet been clarified. On the other hand, some control technologies of SBH for metal/Ge contacts have also been proposed, e.g. the introduction of an insulator interfacial layer, such as Al_2O_3, Ge_2O_3, Ge_3N_4, and SiN [76–78], and the segregation of sulfur at the NiGe/Ge interface [79]. While these technologies provide the ohmic contacts of a metal/n-type Ge interface, the development of metal/Ge contacts is ongoing.

17.5 References

1. Murarka, S. P.: *Silicides for VLSI Applications*, Academic Press, New York (1983).
2. Nicolet, M.-A.; and Lau, S. S.: in *VLSI Electronics, Microstructure Science*, Eispruch, N. and Larrabee, G. (eds), Academic Press, New York, **6**, 329 (1983).
3. Murarka, S. P.: *Intermetallics* **3**, 173–186 (1995).
4. Maex, K.: *Mater. Sci. Eng. Rev.* **R11**, 53–153 (1993).
5. Maex, K.; and van Rossum, M. (eds): *Properties of Metal Silicides*, INSPEC, London (1995).
6. Zhang, S.-L.; and Östling, M.: *Crit. Rev. Solid State Mater. Sci.* **28**(1), 1–129 (2003).
7. Chen, L. J. (ed.): *Silicide Technology for Integrated Circuits*, The Institution of Electrical Engineers (2004).
8. Lavoie, C.; d'Heurle, F. M.; and Zhang, S.-L.: 'Silicides', in *Handbook of Semiconductor Manufacturing Technology*, 2nd edn, Doering, R. and Nishi Y. (eds), CRC Press, Boca Raton, FL (2007).
9. *The International Technology Roadmap for Semiconductor*: http://www.itrs.net/
10. Lau, S. S.; Mayer, J. W.; and Tu, K. N.: *J. Appl. Phys.* **49**, 4005–4010 (1978).
11. van Gurp, G. J.; van der Weg, W. F.; and Sigurd, D.: *J. Appl. Phys.* **49**(7), 4011–4020 (1978).
12. d'Heurle F. M.; and Petersson, C. S.: *Thin Solid Films* **128**, 283–297 (1985).
13. Appelbaum, A. R.; Knoell, V.; and Murarka, S. P.: *J. Appl. Phys.* **57**, 1880–1886 (1985).
14. van Ommen, A. H.; Bulle-Lieuwma, C. W. T.; and Langereis, C.: *J. Appl. Phys.* **64**, 2706–2716 (1988).
15. Wang, Z; Aldrich, D. B.; Chen, Y. L.; Sayers D. E.; and Nemanich, R. J.; *Thin Solid Films* **270**, 555–560 (1995).
16. Aldrich, D. B.; Heck, H. L.; Chen, Y. L.; Sayers, D. E.; and Nemanich, R. J.: *J. Appl. Phys.* **78**, 4958–4965 (1995).

17. Aldrich, D. B.; d'Heurle, F. M.; Sayers, D. E.; and Nemanich, R. J.: *Phys. Rev. B* **53**, 16279–16282 (1996).
18. Aldrich, D. B.; Chen, Y. L.; Sayers, D. E.; Nemanich, R. J.; Ashburn, S. P.; and Öztürk, M. C.: *J. Appl. Phys.* **77**, 5107–5114 (1995).
19. Li, J.; Hong, Q. Z.; Mayer, J. W.; and Rathbun, L.: *J. Appl. Phys.* **67**, 2506–2511 (1990).
20. Zaima, S.; Nakatsuka, O.; Sakai, A.; Murota, J.; and Yasuda, Y.: *Appl. Surf. Sci.* **224**(1–4), 215–221 (2004).
21. Ma, Z.; and Allen, L. H.: *Phys. Rev.* **B49**, 13501–13511 (1994).
22. Mann, R. W.; and Clevenger, L. A.: *J. Electrochem. Soc.* **141**(5), 1347–1350 (1994).
23. Clevenger, L. A.; Harper, J. M. E.; Cabral, C. Jr; Nobili C.; and Ottaviani, G.; Mann, R.: *J. Appl. Phys.* **72**(10), 4978–4980 (1992).
24. van Houtum, H. J. W.; Raaijmakers, I. J. M. M.; and Menting, T. J. M.: *J. Appl. Phys.* **61**(8), 3116–3118 (1987).
25. Lasky, J. B.; Nakos, J. S.; Cain, O. J.; and Geiss, P. J.: *IEEE Trans. Electron Dev.* **38**(2), 262–269 (1991).
26. DiGregorio, J. F.; and Wall, R. N.: *IEEE Trans. Electron Dev.* **47**(2), 313–317 (2000).
27. Ma, Z.; Allen L. H.; and Allman, D. D. J.: *Thin Solid Films* **253**(1–2), 451–455 (1994).
28. Clevenger, L. A.; Roy, R. A.; Cabral, C. Jr; Saenger, K. L.; Brauer, S.; Morales, G.; Ludwig, K. F. Jr; Gifford, G.; Bucchignano, J.; Jordan-Sweet, J.; Dehaven, P.; and Stephenson, G. B.: *J. Mater. Res.* **10**(9), 2355–2359 (1995).
29. Lasky, J. B.; Nakos, J. S.; Cain, O. J.; and Geiss, P. J.: *IEEE Trans. Electron Dev.* **38**(2), 262–269 (1991).
30. Maex, K.; Lauwers, A.; Besser, P.; Kondoh, E.; de Potter, M.; and Steegen, A.: *IEEE Trans. Electron Dev.* **46**(7), 1545–1550 (1999).
31. Ohguro, T.; Saito, M.; Morifuji, E.; Yoshitomi, T.; Morimoto, T. T.; Momose, H. S.; Katsumata,Y.; and Iwai, H.: *IEEE Trans. Electron Dev.* **47**(11), 2208–2213 (2000).
32. Lauwers, A.; de Potter, M.; Chamirian, O.; Lindsay, R.; Demeurisse, C.; Vrancken, C.; and Maex, K.: *Microelectron. Eng.* **60**, 221–230 (2002).
33. Morimoto, T.; Ohguro, T.; Sasaki, H.; Iimura, T.; Kunishima, I.; Suguro, K.; Katakabe, I.; Nakajima, H.; Tsuchiaki, M.; Ono, M.; Katsumata, Y.; and Iwai, H.: *IEEE Trans. Electron Dev.* **42**(5), 915–922 (1995).
34. Lauwers, A.; Steegen, A.; de Potter, M.; Lindsay, R.; Satta, A.; Bender, H.; and Maex, K.: *J. Vac. Sci. Technol.* **B19**, 2026–2037 (2001).
35. Tsuchiya, Y.; Tobioka, A.; Nakatsuka, O.; Ikeda, H.; Sakai, A.; Zaima, S.; and Yasuda, Y.: *Jpn. J. Appl. Phys.* **41**(4B), 2450–2454 (2002).
36. Lavoie, C.; d'Heurle, F. M.; Detavernier, C.; and Cabral, C.: *Microelectron. Eng.* **70**(2–4), 144–157 (2003).
37. Xu, D.-X.; Das, S.R.; Peters, C.J.; Ericleson, L.E.: Thin Solid Films **326**, 143–150 (1998).
38. Okubo, K.; Tsuchiya, Y.; Nakatsuka, O.; Sakai, A.; Zaima, S.; and Yasuda, Y.: *Jpn. J. Appl. Phys.* **43**(4B), 1896–1900 (2004).
39. Mangelinck, D.; Dai, J. Y.; Pan, J. S.; and Lahiri, S. K.: *Appl. Phys. Lett.* **75**(12), 1736–1738 (1999).
40. Liu, J. F.; Chen, H. B.; Feng, J. Y.; and Zhu, J.: *Appl. Phys. Lett.* **77**(14), 2177–2179 (2000).

41. Lee, P. S.; Pey, K. L.; Mangelinck, D.; Ding, J.; Chi, D. Z.; and Chan, L.: *IEEE Electron Dev. Lett.* **22**(12), 568–570 (2001).
42. Tsai, C. J.; Cheng, P. L.; and Yu, K. H.: *Thin Solid Films* **365**(1), 72–76 (2000).
43. Thompson, R. D.; Tu, K. N.; Angillelo, J.; Delage, S.; and Iyer, S. S.: *J. Electrochem. Soc.* **135**(12), 3161–3163 (1988).
44. Luo, J.–S.; Lin, W.-T.; Cheng, C. Y.; and Tsai, W. C.: *J. Appl. Phys.* **82**(7), 3621–3623 (1997).
45. Jarmar, T.; Seger, J.; Ericson, F.; Mangelinck, D.; Smith, U.; and Zhang, S.-L.: *J. Appl. Phys.* **92**(12), 7193–7199 (2002).
46. Nakatsuka, O.; Okubo, K.; Sakai, A.; Ogawa, M.; Yasuda, Y.; and Zaima, S.: *Microelectron. Eng.* **82**(3–4), 479–484 (2005).
47. Cheng, L. W.; Cheng, S. L.; Chen, J. Y.; Chen, L. J.; and Tsui, B. Y.: *Thin Solid Films* **355–356**, 412–416 (1999).
48. Lee, P. S.; Pey, K. L.; Mangelinck, D.; Ding, J.; Wee, A. T. S.; and Chan, L.: *IEEE Electron Dev. Lett.* **21**(11), 568–570 (2000).
49. Choi, C.-J.; Ok, Y.-W.; Hullavarad, S. S.; Seong, T.-Y.; Lee, K.-M.; Lee, J.-H.; and Park, Y.-J.: *J. Electrochem. Soc.* **149**(9), G517–G521 (2002).
50. Thompson, R. D.; Tu, K. N.; Angillelo, J.; Delage, S.; and Iyer, S. S.: *J. Electrochem.* Soc. **135**(12), 3161–3163 (1988).
51. Detaverniera, C.; and Lavoie, C.: *Appl. Phys. Lett.* **84**, 3549–3551 (2004).
52. Jin, L. J.; Pey, K. L.; Choi, W. K.; Fitzgerald, E. A.; Antoniadis, D. A.; Pitera, A. J.; Lee, M. L.; Chi, D. Z.; Rahman, M.A.; Osipowicz, T.; and Tung, C. H.: *J. Appl. Phys.* **98**, 033520 (2005).
53. Lauwers, A; van Dal, P.; Verheyena, M. J. H.; Chamirian, O.; Demeurisse, C.; Mertens, S.; Vrancken, C.; Verheyden, K.; Funk, K.; and Kittl, J. A.: *Microelectron. Eng.* **83**, 2268–2271 (2006).
54. Yu, A. Y. C.: *Solid-State Electron.* **13**, 239–247 (1970).
55. Chang, C. Y.; Fang, Y. K.; and Sze, S. M.: *Solid-State Electron.* **14**, 541–550 (1971).
56. Iwai, H.; Ohguro, T.; and Ohmi, S., *Microelectron. Eng.* **60**, 157 (2002).
57. Goto, K.-I.; Fushida, A.; Watanabe, J.; Sukegawa, T.; Tada, Y.; Nakamura, T.; Yamazaki, T.; and Sugii, T.: *IEEE Trans. Electron Dev.* **46**(1), 117–124 (1999).
58. Ikegami, H.; Ikeda, H.; Zaima, S.; and Yasuda, Y.: *Appl. Surf. Sci.* **117–118**(2), 275–279 (1997).
59. Tsuchiaki, M.; Hongo, C.; Takashima, A.; and Ohuchi, K.: *Jpn. J. Appl. Phys.* **41**, 2437 (2002).
60. Tsuchiaki, M.; Ohuchi, K.; and Hongo, C.: *Jpn. J. Appl. Phys.* **43**, 5166 (2004).
61. Tsuchiaki, M.; Murakoshi, A.; and Hongo, C.: *Jpn. J. Appl. Phys.* **42**, 1847 (2003).
62. Tsuchiaki, M.; Ohuchi, A.; and Nishiyama A.: *Jpn. J. Appl. Phys.* **44**, 1673 (2005).
63. van de Walle, C. G.; and Martin, R.: *Phys. Rev. B* **34**(8), 5621–5634 (1986).
64. Zaima, S.; and Yasuda, Y.: *J. Vac. Sci. Technol.* **B16**(5), 2623–2628 (1998).
65. Ghani, T.; Armstrong, M.; Auth, C.; Bost, M.; Charvat, P.; Glass, G.; Hoffmann, T.; Johnson, K.; Kenyon, C.; Klaus, J.; McIntyre, B.; Mistry, K.; Murthy, A.; Sandford, J.; Silberstein, M.; Sivakumar, S.; Smith, P.; Zawadzki, K.; Thompson, S.; and Bohr, M.: *IEDM Tech. Dig.*, 978–980 (2003).
66. Ang, K. W.; King, J. C.; Bliznetsov, V.; Balasubramanian, A. D.; Li, M. F.; Samudra, G.; and Yeo, Y. C.: *IEDM Tech. Dig.*, 1069–1071 (2004).

67. Sze, S. M.; and Ng, K.K.: *Physics of Semiconductor Devices*, 3rd edn, Wiley, New York, 2007.
68. Qin, M.; Poon, V. M. C.; and Hoa, S. C. H.: *J. Electrochem. Soc.* **148**(5), G271–G274 (2001).
69. Maszara, W. P.: *J. Electrochem. Soc.*, **152**(7), G550–G555 (2005).
70. Sim, J. H.; Wen, H. C.; Lu, J. P.; and Kwong, D. L.: *IEEE Trans. Electron Dev. Lett.* **24**(10), 631–633 (2003).
71. Kittl, J. A.; Lauwers, A.; Pawlak, M. A.; van Dal, M. J. H.; Veloso, A.; Anil, K. G.; Pourtois, G.; Demeurisse, C.; Schram, T.; Brijs, B.; de Potter, M.; Vrancken, C.; and Maex, K.: *Microelectron. Eng.* **82**(3–4), 441–448 (2005).
72. Ransom, C. M., Jackson, T. N., and DeGelormo, J. F.: *IEEE Trans. Electron Dev.* **38**, 2695 (1991).
73. Claeys, C. L.; and Simoen, E.: *Germanium-based Technologies: From Materials to Devices*, Elsevier, New York 2007.
74. Dimoulas, A.; Tsipas, P.; and Sotiropoulos, A.: *Appl. Phys. Lett.* **89**, 252110 (2006).
75. Nishimura, T.; Kita K.; and Toriumi, A.: *Appl. Phys. Lett.* **91**, 123123 (2007).
76. Nishimura, T.; Kita, K; and Toriumi, A.: *Appl. Phys. Exp.* **1**, 051406 (2008).
77. Lieten, R. R.; Degroote, S.; Kuijk, M.; and Borghs, G.: *Appl. Phys. Lett.* **92**, 022106 (2008).
78. Kobayashi, M.; Kinoshita, A.; Saraswat, K.; Wong, H.-S. P.; and Nishi, Y.: *J. Appl. Phys.* **105**, 023702 (2009).
79. Ikeda, K.; Yamashita, Y.; Sugiyama, N.; Taoka, N.; and Takagi, S.: *Appl. Phys. Lett.* **88**, 152115 (2006).

18

Silicon–germanium (SiGe) heterojunction bipolar transistor (HBT) and bipolar complementary metal oxide semiconductor (BiCMOS) technologies

K. WASHIO, Hitachi Ltd, Japan

Abstract: Silicon–germanium (SiGe) heterojunction bipolar transistor (HBT) and bipolar CMOS (BiCMOS) technologies are reviewed in terms of their current status and potential future directions. SiGe HBTs and BiCMOS are promising candidates for both high-speed digital operation and high-frequency analog operation for ICs and LSIs in multigigabit data-communication systems and wide-bandwidth wireless communication systems. SiGe epitaxial growth technologies, typical device structures, fabrication processes, and transistor characteristics of SiGe HBTs and BiCMOS are described. ICs and LSIs based on SiGe HBTs and BiCMOS for optical fiber links and wireless communications are also described.

Key words: silicon–germanium (SiGe), heterojunction bipolar transistor (HBT), bipolar CMOS (BiCMOS), data communication, wireless communication.

18.1 Introduction

Devices achieving both high-speed digital operation with sophisticated functions and high-frequency analog operation are the key components in ICs and LSIs for multigigabit data-communication systems and wide-bandwidth wireless communication systems. To meet the sustained growing demands for such systems, high-speed and high-frequency performance of SiGe HBT and BiCMOS devices have been improved.[1] SiGe HBT and BiCMOS technologies are promising because they can be fabricated by the well-established Si process, which is fully compatible with the CMOS process, and they can provide high fabrication yield and integration of high-quality passive elements for the widespread applications in high-speed digital and RF analog ICs/LSIs. Innovative SiGe hetero-epitaxial-growth technologies, ingenious device structures, and novel circuit schemes have been investigated; as a result, both high-speed and high-frequency performances have been attained.

This chapter is organized as follows. SiGe epitaxial growth technology – a key process technology to fabricate SiGe HBTs – is described in Section 18.2. Typical device structures, fabrication processes, and transistor characteristics of SiGe HBTs are described in Section 18.3. Typical device structures, fabrication processes, and device characteristics concerning SiGe BiCMOS technologies are presented in Section 18.4. As applications of the SiGe HBT and BiCMOS technologies, ICs/LSIs for optical fiber links and wireless communications are described in Section 18.5. Finally, the current status and potential directions for the future are summarized in Section 18.6.

18.2 Epitaxial growth

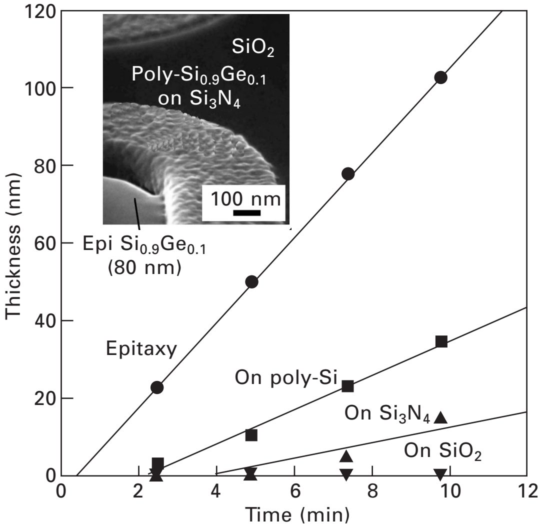
In this section, epitaxial growth techniques for fabricating SiGe HBTs, an effect of carbon incorporation to reduce boron diffusivity, and a continuous epitaxial growth of Si/SiGe multilayers are described.

18.2.1 SiGe epitaxial growth techniques

To fabricate a SiGe HBT, a strained SiGe base layer is grown on a Si collector layer. Epitaxial growth techniques can be classified roughly into two kinds according to the deposition area, namely blanket or selective epitaxial growth. In the case of blanket epitaxial growth, SiGe hetero-epitaxial layers are deposited on the whole wafer area. On the other hand, in the case of selective epitaxial growth (SEG), SiGe layers are formed only on the Si surface exposed by patterning. The choice of epitaxial growth technique therefore strongly affects process sequences and device structures of SiGe HBTs, as described in the next section.

There are a variety of epitaxial growth techniques and many related properties, but their details are left to other papers or books owing to the limitations of space here. The following section focuses on the critical thickness of a SiGe epitaxially grown layer. This is because misfit dislocations significantly degrade transistor performance such as leakage current between the collector and emitter. In the Si/SiGe/Si heterostructure used in a SiGe HBT, the critical thickness is given by extending the basic image-force equilibrium model.[2] The critical thickness decreases with Ge content but slightly increases with Si-cap layer thickness. At typical Ge content in the SiGe base layer of 10 to 20%, the critical thickness is in the range of 60 to 30 nm. SiGe layer thickness and Ge content should therefore be carefully designed to avoid forming defects caused by stress relaxation. The phenomenon related to the window size of epitaxial growth is briefly described in the following. As the window size is reduced to less than about 10 μm, dislocation density in the SiGe/Si heterostructure decreases, likely owing to the blocking of misfit-dislocation extension and the edge-induced strain relaxation.[3] SEG is therefore more preferable to achieve a stable and high Ge-content SiGe epitaxial layer, because the size of the growth window in SEG is usually in the order of microns.

SEG is advantageous because it makes it possible to create a self-aligned SiGe HBT process. However, SEG has to be done on patterned wafers, and its growth conditions are sensitive to the nature of the wafer surface. This is because selectivity is obtained by precisely controlling the nucleation step for both the Si and dielectric surfaces under the balanced gas flows used for deposition and etching. However, especially when applying SEG to form SiGe HBTs, it is not necessary to obtain perfect SEG. Namely, it is sufficient not to deposit any layers on the dielectric surfaces during growth of the SiGe layers on Si. This so-called 'quasi SEG' can be achieved by utilizing the difference between incubation times in accordance with surface materials.^{4} The growth time dependence of SiGe layer thickness on Si, poly-Si, Si_3N_4, and SiO_2 surfaces for this method with no etching gas flow is shown in Fig. 18.1. The incubation times in the cases of growth on poly-Si and Si_3N_4 are longer than that on Si. Moreover, no layer is deposited on SiO_2 even at the growth time for a 100-nm-thick SiGe layer on Si. This means that SEG can be attained by using a SiO_2 mask when the SiGe layer thickness is limited in this manner. This quasi SEG has less sensitivity to the nature of the patterned wafer because it involves no etching gas; that is, loading effect (variable deposition rates and epitaxial layer compositions over the wafer which are related to variations in surface topology) and facet formation are very small, and the thickness



18.1 Thickness of $Si_{0.85}Ge_{0.15}$ grown on Si(100) substrate, poly-Si, Si_3N_4, and SiO_2 surfaces, as a function of growth time. The inset is an SEM bird's-eye view of SiGe selective growth when an epitaxy 80-nm-thick $Si_{0.9}Ge_{0.1}$ layer is grown selectively with respect to only SiO_2.

of the epitaxial growth layer is quite uniform. It can therefore be applied in existing LSI fabrication processes.

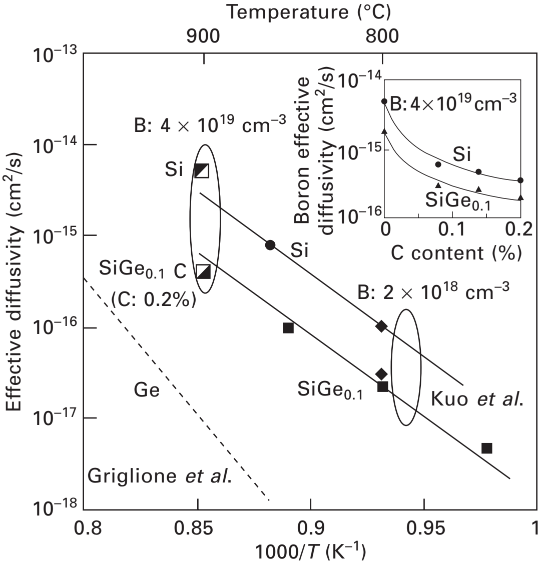
18.2.2 Carbon-doped SiGe growth

It is well known that thinning of the base width of SiGe HBTs significantly improves typical characteristics that indicate high-speed and high-frequency performance, that is, cutoff frequency and maximum oscillation frequency. To enhance operating speed and frequency of SiGe HBTs, impurity-profile engineering, especially for boron (B) in the base, is therefore very important.

It is known that B effective diffusivity in $Si_{0.9}Ge_{0.1}$ is about one-third of that in Si.[5] However, when high B concentration for a thin base layer is applied to decrease base sheet resistance, B effective diffusivity is increased by transient enhanced diffusion (the dopant diffusion is typically enhanced by several orders of magnitude in high-doped regions). To improve high-speed performance and characteristic controllability, it is effective to incorporate substitutional carbon (C) in the base layer and reduce B diffusivity, as shown in Fig. 18.2.[6,7] This effect is explained by C-induced undersaturation of Si interstitials that is caused by replacement of interstitial Si by substitutional C. That is, B at a substitutional site is easily replaced by interstitial Si, and its diffusivity is enhanced; however, in the case of C incorporation, B tends to remain at a substitutional site because of lower concentration of Si interstitials, and its diffusivity is decreased. By incorporating 0.2% C, B effective diffusivity at a B concentration of 4×10^{19} cm^{-3} decreases by less than one order of magnitude (see inset of Fig. 18.2), and that in $Si_{0.898}Ge_{0.1}C_{0.002}$ at a B concentration of 4×10^{19} cm^{-3} is even lower than that in $Si_{0.9}Ge_{0.1}$ at a B concentration of 2×10^{18} cm^{-3}. It should be noted here that Ge diffusivity is much lower than B diffusivity,[8] so the Ge profile is almost kept as-grown after the state-of-the-art low-temperature fabrication process. Heterostructure-related characteristics of SiGe HBT are therefore determined by the initial Ge profile.

In regard to $Si_{1-x-y}Ge_xC_y$ epitaxial growth, surface roughness related to Ge and C contents should be considered carefully in the case of high C content.[7] The surface roughness for low C content (i.e., less than 0.4%) is almost constant. However, for higher C content, the surface roughness depends on the ratio of Ge to C contents. In particular, the surface roughness reaches a minimum at a ratio of about 12, at which the lattice strain is close to being fully alleviated. Incorporation of C at a ratio deviating from this condition causes large surface roughness due to local lattice strain.

Last but not least on C incorporation is its effect in alleviating compressive lattice strain in the SiGe layer. That is, C causes a strain-compensation effect because it induces tensile strain. The critical thickness of the SiGe layer



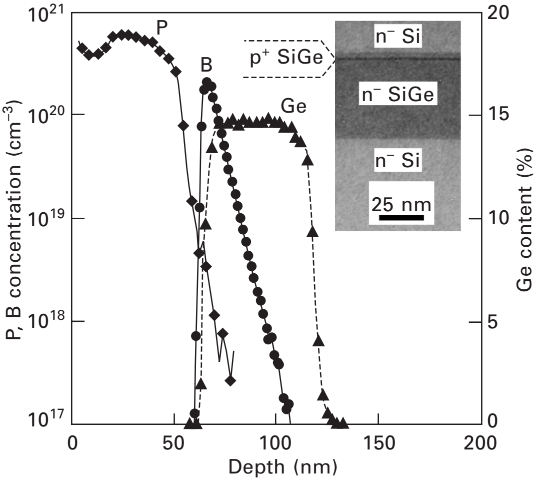
18.2 Effective diffusivities for high-concentration B diffusion in Si, $SiGe_{0.1}$, and $SiGe_{0.1}C$. The inset shows dependence of B effective diffusivity on C content for both Si and $SiGe_{0.1}$.

can therefore be increased by incorporating C. Accompanying this effect, C incorporation increases the bandgap in the SiGe system. However, the incorporated C content is much less than the Ge content, so the effective widening of the bandgap is small (i.e., about 10 meV). The smaller lattice strain in the SiGeC system has a beneficial effect on HBT performance. In other words, in the case of SiGe HBTs with different-sized emitters, a slight change of the bandgap due to the change of effective strain is usually observed. (The bandgap increases with emitter width.) However, in the case of SiGeC HBTs, the bandgap remains almost constant. This effect makes it easy to extract HBT model parameters and to use HBTs with a variety of emitter sizes, which depend on circuit design.

18.2.3 Continuous growth of Si/SiGe multilayer

To apply SiGe HBTs to millimeter-wave wireless communications with high-speed data transmission capability, ultra-high-frequency performance must be achieved. To satisfy this requirement, base width should be minimized, and doping concentration should be maximized. Fabricating a Si/SiGe multilayer to compose an HBT's emitter, base, and collector in one epitaxial growth step is therefore considered to be a promising way. Because one-step epitaxial growth can be done under low temperature, it produces a narrow and abrupt base profile. Furthermore, it provides high doping concentration with good crystallinity. It should be noted here that the Si/SiGe multilayer has to be formed with different types of dopant in the SiGe HBT, namely n-type for the emitter and collector, and p-type for the base.

Continuous-epitaxial-growth technology – based on the cold-wall rapid thermal (RT) ultrahigh vacuum/chemical vapor deposition (UHV/CVD) system – is one solution for satisfying the above-mentioned ultra-high-frequency performance requirement.[9] Since this system utilizes a UHV transfer chamber (background pressure below 10^{-7} Pa) and two growth chambers for n-type and p-type doping, contamination during wafer transfer to each chamber is minimized without the need for a high-temperature cleaning process. Furthermore, the system can operate in both UHV/CVD mode (deposition pressure below 1 Pa) and LPCVD mode. Ge and B profiles can be precisely controlled by forming a SiGe base layer in UHV/CVD mode, and good crystal quality and highly P-doped collector and emitter layers can be obtained in LPCVD mode. The SIMS depth profiles of B and P concentrations, and Ge content, in a SiGe HBT are shown in Fig. 18.3. A very thin base (3.5 nm thick) with B doping concentration of more than 2×10^{20} cm^{-3} is successfully formed (the tail of the B profile is caused by SIMS analysis), and other impurity profiles (i.e., P and Ge) are precisely controlled. A cross-sectional TEM view of the intrinsic n$^-$-Si/p$^+$-SiGe/n$^-$-Si multilayer (see inset of Fig. 18.3) confirms good crystal quality. This result demonstrates that continuous epitaxial growth without high-temperature processing is a promising way to dramatically improve high-frequency performance of SiGe HBTs.[10]



18.3 SIMS depth profiles of dopant concentrations and Ge content in a SiGe HBT. The inset is a cross-sectional TEM view of the intrinsic n$^-$-Si/p$^+$-SiGe/n$^-$-Si/n$^+$-Si multilayer.

18.3 Silicon–germanium (SiGe) heterojunction bipolar transistor (HBT)

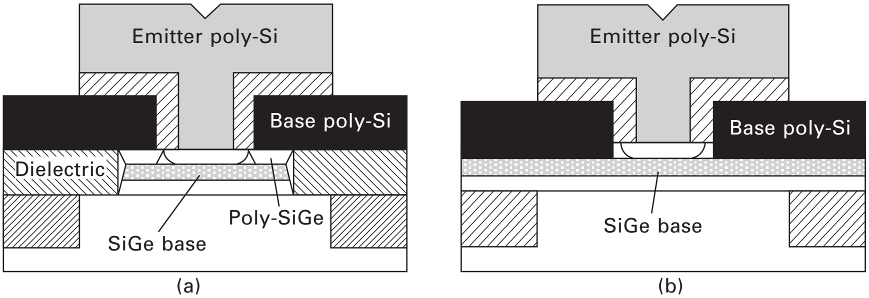
In this section, device structures of SiGe HBTs for obtaining high-speed and high-frequency performance, a fabrication process to form self-aligned SiGe HBTs, and transistor characteristics are described.

18.3.1 Device structure

To bring out the inherent high-speed and high-frequency characteristics of a SiGe HBT, its device structure has to be ideally constructed only with essential parts that contribute to transistor action. In other words, parasitic areas, which are used to connect to intrinsic emitter, base, and collector regions, should be minimized. These parasitic areas mainly arise from the margin for mask alignment in the lithography process steps. It is therefore very important to apply a self-aligned manner to define the positions of each intrinsic region. Accordingly, SiGe HBTs with various self-aligned device structures to reduce parasitic capacitances and resistances have been developed. They can be roughly categorized into two types. Schematic cross-sectional views of the two types of self-aligned SiGe HBT are shown in Fig. 18.4. Although they appear similar at a glance, their process sequences are quite different, as described in the next section. The differences in their device structures come from the SiGe epitaxial growth technologies used, namely, selective epitaxial growth is used to form the intrinsic SiGe base region in a self-aligned manner, called a SEG-SA structure, shown in Fig. 18.4(a),[11,12] and blanket epitaxial growth and a sacrificial-emitter process (SEP) sequence is applied to form the self-aligned structure in Fig. 18.4(b),[13,14] namely a SEP-SA structure. In both SiGe HBTs, these self-aligned device structures reduce the size of parasitic areas in common. The distance for isolating the base and emitter poly-Si electrodes can be minimized by defining the thickness of the dielectric layer. Base resistance generated at the link base (the region between the periphery of emitter and poly-Si base electrode) is therefore reduced. Furthermore, especially in the SEG-SA structure, a collector-base junction region is also defined from the emitter region in a self-aligned manner, and almost all of the base poly-Si electrode is on the thick dielectric layer. Collector-base junction capacitance is thus greatly decreased. As a result of applying the above-mentioned self-aligned device structures, high-speed and high-frequency performance can be achieved in conjunction with an intrinsic high-potential SiGe heterostructure.

18.3.2 Fabrication process

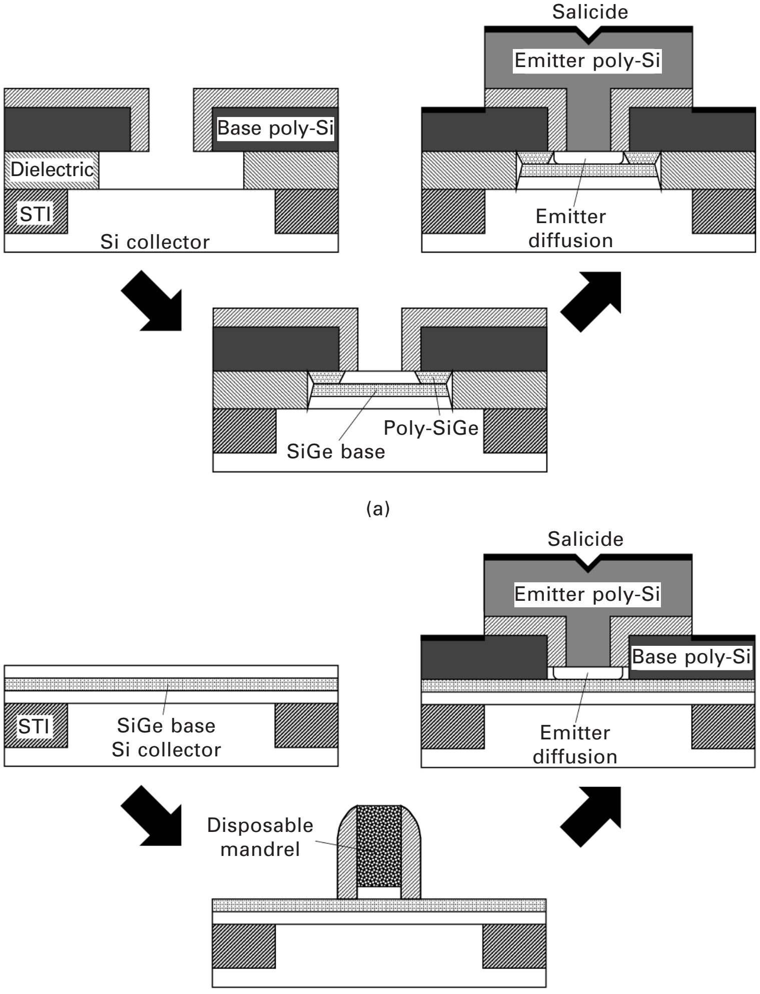
Even though the two kinds of device structures, SEG-SA and SEP-SA, are very similar, their fabrication process sequences are different, as shown in



18.4 Schematic cross-sectional views of self-aligned SiGe HBTs. Selective epitaxial growth in (a) and blanket epitaxial growth with sacrificial-emitter process sequence in (b) are used. (From Ref. 27, with permission.)

Figs 18.5(a) for SEG-SA and (b) for SEP-SA. In the SEG-SA SiGe HBT, a cavity enclosed by a base poly-Si electrode and dielectric materials is preformed. This cavity is formed by patterning of the base poly-Si electrode and subsequent etching of dielectric layers. Following that, the cavity is filled up by a SiGe multilayer, consisting of a Si cap, a SiGe base, and a lightly doped SiGe collector, by selective epitaxial growth. Here, a poly-SiGe layer is simultaneously formed as an extrinsic base for providing the link between the intrinsic SiGe base and the base poly-Si electrode. This link region can therefore be made small because no mask alignment margin is necessary. This is a very effective way to reduce parasitic collector capacitance. After that, an *in situ* P-doped poly-Si (IDP) layer is deposited for an emitter electrode, and an intrinsic emitter is formed by diffusion from the IDP layer into the Si cap layer. Finally, metal-silicide (such as $TiSi_2$ and $CoSi_2$) layers are formed simultaneously on all electrodes in a self-aligned manner.

On the other hand, in the SEP-SA SiGe HBT, the SiGe multilayer is grown on the whole surface and a disposable mandrel is patterned to define the emitter region. By a process similar to a gate patterning of a MOS transistor, the dielectric layer for separating the emitter and base regions is formed at the sidewall of the disposable mandrel. Following that, a poly-Si or a single-Si layer is formed as the base electrode by a planarization process. As a result, the base poly-Si electrode is formed self-aligned to the emitter. Subsequent process steps are the same as those in the case of the SEG-SA process sequence. In the SEP-SA structure as well as in the SEG-SA structure, a narrow emitter–base separation can be formed. The width of the link region between the intrinsic SiGe base and the base poly-Si electrode might be slightly wider than that in the SEG-SA structure due to the mask alignment. This causes a slight increase in collector capacitance compared with that in the case of the SEG-SA structure. However, recent lithography processes provide a very small alignment margin as well as fine pattern


Salicide
Emitter poly-Si
Base poly-Si
Dielectric
STI
Si collector
Emitter
diffusion
Poly-SiGe
SiGe base
Salicide
Emitter poly-Si
Base poly-Si
STI
SiGe base
Si collector
Emitter
diffusion
Disposable
mandrel

(a)

(b)

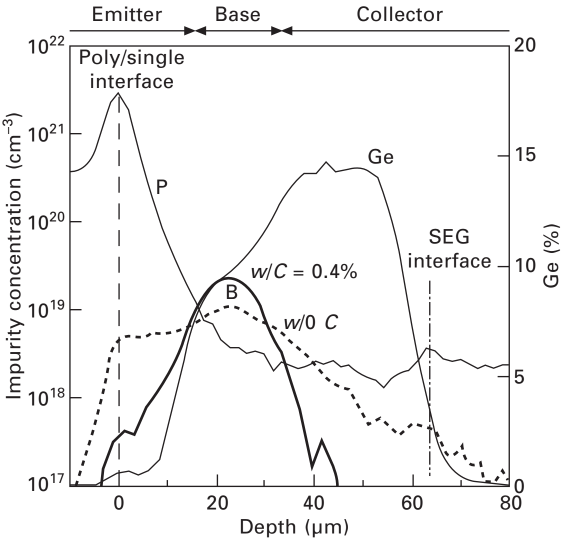
18.5 Fabrication process sequences of SiGe HBTs with self-aligned structures by using (a) elective-epitaxial-growth and (b) sacrificial-emitter-process.

formation, so the difference between the SEG-SA and SEP-SA structures is becoming smaller, and less attention will be paid in future to the self-aligned structure itself for cutting-edge transistors.

To improve high-frequency performance, scaling of emitter width is essential.[15] This is because reducing emitter width allows intrinsic base resistance to be decreased; however, doing so increases emitter resistance. Consequently, for narrow-emitter SiGe HBTs, the conventional poly-Si emitter electrode has been replaced by a heavily doped single-crystal Si emitter layer,[16] which is also formed by epitaxial growth. There is therefore no interfacial native oxide layer between the mono-Si emitter electrode and the intrinsic emitter. As a result, the emitter resistance is reduced. Even so, the epitaxial growth conditions should be carefully optimized. Since the intrinsic SiGe base has already formed before this emitter epitaxial growth is performed, the thermal budget of the pre-cleaning process must be reduced to keep the base width thin. Furthermore, to decrease resistivity of the mono-Si emitter layer, sufficiently high-concentration doping should be carried out, obtaining good crystal quality, by optimizing the growth temperature and pressure.[17] Applying the mono-Si emitter layer, with no interfacial oxide layer between the emitter electrode and intrinsic emitter, brings the beneficial effect of reducing low-frequency noise (1/f noise)[18] – a key factor in microwave analog circuits because it generates undesired noise. Accordingly, the mono-Si emitter also improves RF circuit performance.

18.3.3 Transistor characteristics

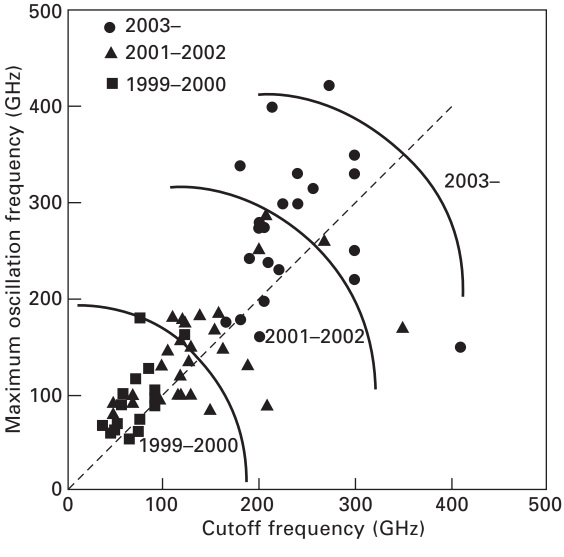
The design of the Ge profile primarily determines SiGe HBT characteristics. To control SiGe HBT characteristics precisely, the Ge profile has therefore been investigated and optimized.[19,20] In these studies, it was found that a Ge profile featuring dual plateaus with an intermediately graded slope (DPIG) in the base, as shown in Fig. 18.6, can tolerate fluctuation of emitter-base



18.6 Intrinsic impurity profiles for SiGe and SiGeC HBTs. Ge profile featuring dual plateaus with intermediate slope in the base layer is used to achieve high-precision SiGe HBT characteristics. (From Ref. 27, with permission.)

junction position. This Ge profile is composed as follows: two plateaus (one with Ge contents of 10% on the side towards the surface and one with that of 15% on the side towards the substrate), a 5-nm-thick Ge-graded layer between the plateaus, and two Ge-graded layers at the interfaces with the SEG and Si-cap layers. This DPIG Ge profile enables precise controllability of collector current against deviation in the position of the emitter–base junction. The intermediate Ge-graded layer provides an internal drift field in the conduction band and enables a short base transit time and a high early voltage. Figure 18.6 shows that B diffusivity can be suppressed by C incorporation (as described in Section 18.2.2) in the intrinsic impurity profile of the SiGeC HBT.

Thanks to efforts to minimize size of extrinsic regions in the self-aligned SiGe HBTs, benchmarks for both high-speed digital operation and high-frequency analog performance, i.e., cutoff frequency f_T and maximum oscillation frequency f_{max}, respectively, have been improved simultaneously, as shown in Fig. 18.7. In particular, both f_T and f_{max} exceeded 100 GHz in 2000[21] and 250 GHz in 2002.[22] This rapid improvement in high-speed and high-frequency performance is attributed to thinning the base width combined with an ingenious device structure. However, despite the guidelines laid down in the International Technology Roadmap for Semiconductors (ITRS) (which advocates improving both f_T and f_{max}[23]), after 2003, more emphasis has been placed on f_{max}.[24] This direction can be considered appropriate because emerging applications, such as millimeter-wave or terahertz wireless

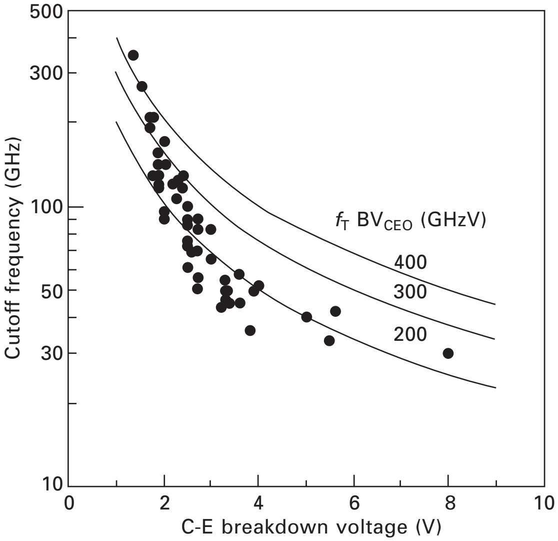


18.7 Improvement in cutoff frequency and maximum oscillation frequency from 1999 to 2009.

systems, require high RF performance and $f_{\max}$ is the main figure of merit (FOM) for RF circuits. In detail, due to a trade-off between base width and base resistance (that is, a trade-off between f_T and $f_{\max}$), the direction of future research tends to target achieving higher $f_{\max}$ while keeping f_T at about 300 GHz. This trend is reasonable in the sense that it is beneficial from the scaling viewpoint by using fine lithography in state-of-the-art CMOS processes and in the sense that it gives sufficient breakdown voltage to obtain wide dynamic range and high output power.

From the standpoint of circuit design, breakdown voltage as well as high-speed performance is a very important characteristic. When breakdown voltage is considered, the well-known trade-off relationship known as the Johnson limit comes into play.[25] The dependence of f_T on collector–emitter breakdown voltage BV_{CEO} is shown in Fig. 18.8. Advanced device structures raise the $f_T BV_{CEO}$ product from 200 GHzV to more than 400 GHzV in the high-f_T region. However, BV_{CEO} drops to less than 1.5 V for f_T of more than 300 GHz. Furthermore, high f_T values are usually achieved at high current density, which is undesirable from the viewpoint of the self-heating problem. An emphasis on $f_{\max}$ might be a better direction for future research.

Integration of SiGe HBTs with different breakdown voltages is important because a high-breakdown-voltage SiGe HBT is necessary to handle high voltage in a driver circuit or an I/O circuit. The balance between breakdown voltage and high-frequency performance is usually controlled by adjusting the



18.8 Dependence of cutoff frequency f_T on collector–emitter breakdown voltage BV_{CEO}. Advanced device structure raises $f_T BV_{CEO}$ product to more than 400 GHzV in the high f_T-region. (From Ref. 27, with permission.)

collector doping concentration through methods such as selective phosphorus implantation with an additional mask step.[19,26]

The FOM for high-speed digital circuits is the gate delay of the logic circuit. ECL (emitter-coupled logic) and CML (current mode logic) gate delays have tended to halve every five years since the early 1990s,[27] and gate delay has presently reached less than 3 ps.[28] This means that a simple logic circuit like a frequency divider has the potential to operate up to about 200 GHz. It can therefore be considered that sufficient high-speed digital performance for many kinds of communication systems has already been achieved. However, it should be added that further reduction of gate delay below 2 ps might be very difficult because of the deterioration in the balance between intrinsic speed and parasitic factors caused by scaling or the f_{max}-emphasized research direction as described previously.

In contrast to the previously described approaches focusing on high-frequency or high-speed performance, another approach focuses on reducing collector current density J_{C} while keeping appropriate f_{T} for the micro- to millimeter wavebands.[29] This approach corresponds to the growing demand for low-power solutions that is accompanying the spread of mobile communications. It is well known that there is a trade-off relationship between J_{C} and f_{T}. It is therefore not easy to achieve both high f_{T} and low J_{C} simultaneously. To overcome this trade-off, an abrupt emitter–base junction formed by continuous epitaxial growth, as described in Section 18.2.3, and a low-temperature fabrication process have been implemented. By optimizing the thickness of the n^- Si spacer in the emitter to be fully depleted while the HBT is operating, emitter–base junction capacitance can be effectively reduced while adequate base transit time is maintained. As a result, ultra-low-power performance with peak f_{T} of up to 200 GHz is achieved.

There are two more worthy research topics that aim to improve circuit performance and expand the application fields of SiGe HBTs. One is on pnp transistors and the other is on high-voltage transistors. To achieve high linearity and low power consumption, complementary circuit schemes are preferable. Accordingly, integration of pnp transistors with npn SiGe HBTs is being investigated.[30] To achieve well-matched performance for both transistors, a heterostructure is also applied to pnp transistors; however, this approach tends to increase process complexity and cost. Moreover, these simplified integration schemes partly limit performance. However, recent dedicated works are overcoming this dilemma.[31–33]

Integration of a transistor that can block high voltages while keeping its high-frequency performance is required. To satisfy this requirement, a novel vertical trench SiGe HBT is proposed.[34] A concept called the reduced surface field (RESURF) is applied to reshape the electric field distribution in the collector. A vertical trench containing a field plate and a linearly graded doping profile in the collector drift region provide a uniform electric field.

According to simulation results on this SiGe HBT, the $f_T BV_{CBO}$ product was improved to about 2400 GHzV.

18.4 Silicon–germanium (SiGe) bipolar complementary metal oxide semiconductors (BiCMOS)

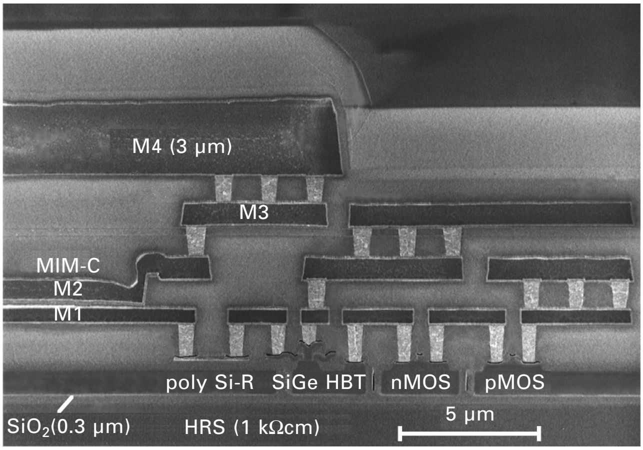
In this section, suitable device structures for integrating scaled CMOS with high-speed SiGe HBTs, fabrication process schemes for applying scaled CMOS to SiGe BiCMOS, and device characteristics of SiGe BiCMOS are described.

18.4.1 Device structure

To meet the demand for multigigabit optical fiber link systems and wide-bandwidth wireless communications, both high-frequency analog operation and high-speed digital operation with sophisticated functions should be implemented simultaneously. From this point of view, integration of high-frequency SiGe HBTs with high-density CMOS is one of the most promising candidates. The basic concept of SiGe BiCMOS technology involves integrating individual transistors on the same chip as they are. Device structures of SiGe HBTs and CMOS are therefore designed in the same manner as those fabricated separately. Furthermore, their device characteristics must remain unchanged. These requirements are satisfied by improving the SiGe BiCMOS fabrication process as described in the next section.

A cross-sectional SEM image of 0.2-μm SiGe BiCMOS, as an example of the many kinds of early SiGe BiCMOS technologies, is shown in Fig. 18.9.[1] In this example, self-aligned SiGe HBTs and CMOS are integrated on a Si-on-insulator (SOI) wafer. (Note that the SOI wafer is not essential, and integrated passive devices are described in Section 18.4.3.) As in the figure, the device structures are kept as individually fabricated.

Some novel integration schemes, in which the device structure of the SiGe HBT is modified, have recently been developed. This is because CMOS fabricated on thin SOI has become widely used as an advanced technology. A SOI wafer has advantages such as complete isolation, latch-up immunity, weak substrate coupling, and enhanced passive device performance. However, the problem of incompatibility of high-performance SiGe HBTs arises with thin SOI CMOS. This is because a typical SiGe HBT is a vertically acting device, so a certain degree of Si thickness (typically about 0.2 μm at lowest) is necessary to fabricate it. A thin SOI is, however, only 0.1 μm thick or less. To overcome this hurdle, two unique approaches have been taken.[35,36] One is to draw the collector drift path laterally, and the other is to form the collector regions in the Si substrate below the buried oxide. There are pros



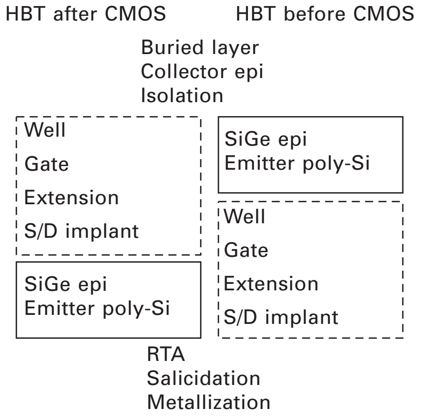
18.9 Cross-sectional SEM image of 0.2-µm self-aligned SEG SiGe HBT, CMOS transistor, and poly-Si resistor on 200-mm SOI (1-µm-thick Si on 0.3-µm-thick insulator) wafer based on high-resistivity substrate. Four-level metal (M1–M4) layer structure is used for interconnection. (From Ref. 1, with permission.)

and cons to both approaches, but they have both been successfully applied to thin SOI CMOS technology.

18.4.2 Fabrication process

As for the SiGe BiCMOS fabrication process, most attention has been paid to ensuring that the performance of neither SiGe HBT nor CMOS degrades. The basic process sequences for each device are already established, so they can be separated into common parts and particular parts. The SiGe BiCMOS fabrication process can then be built up as a modular scheme. Roughly speaking, there are two kinds of schemes for integrating scaled CMOS with SiGe NBT, namely HBT-after-CMOS[37] and HBT-before-CMOS,[38] as shown in Fig. 18.10.

To form the buried layers, lightly doped collectors, and isolations, a standard process is carried out. In the case of the HBT-after-CMOS scheme, the CMOS process module, which includes the formation of the well regions, gate patterns, source/drain extension areas, and source/drain diffusion layers, is performed before SiGe epitaxial growth. In this process sequence, it is very important to reduce the thermal budget during SiGe epitaxy so as not to affect the ready-formed CMOS characteristics. A low-temperature (about 700°C or lower) pre-cleaning process for removing contaminants on the Si surface, such as oxygen and carbon (which degrade the quality of the epitaxial layer), is therefore very important. In the case of the HBT-before-CMOS scheme, the SiGe HBT process module, which includes SiGe epitaxial



18.10 Integration process schemes of SiGe BiCMOS for high-speed/high-frequency HBT embedded with scaled CMOS. (From Ref. 27, with permission.)

growth and formation of an emitter electrode, is inserted before the CMOS process module. In this process sequence, it is essential to maintain the base impurity profile during the CMOS thermal steps so as not to degrade SiGe HBT performance. C-doped SiGe epitaxial growth to suppress B out-diffusion, as described in Section 18.2.2, is therefore extensively adopted.

One of the most important aspects in the fabrication process of SiGe BiCMOS is reducing cost. Simply combining the fabrication process modules of SiGe HBT and CMOS technologies increases the number of process sequences and thus significantly adds complexity. To apply SiGe BiCMOS technology to widespread use or high-volume applications, it is therefore very important to establish a low-cost process. Several approaches can be taken to achieve this goal. They include a simplified collector construction without an epitaxially buried subcollector or deep trench isolation,[39] replacement of HBT's specialized processes with alternatives like sharing implants with the CMOS process flow, and a one-mask HBT module in which a single mask only is used for the HBT.[40] In all these process schemes, most attention is paid to minimizing the degradation of SiGe HBT performance while keeping complete compatibility with the base CMOS process.

The most attractive feature of SiGe BiCMOS is its functionality. The most popular way to improve the functionality is by integrating high-frequency SiGe HBT analog circuits with high-density CMOS digital circuits. To achieve this integration, a scaled CMOS module – in which process conditions such as gate shape and thermal budget are precisely controlled to form a narrow channel region and a shallow diffusion layer – is applied. As for this CMOS module, a trend in CMOS gate scaling concerning the SiGe BiCMOS fabrication process is touched on in the following. To achieve sophisticated functions with high-speed SiGe HBT front-end circuits, large-scale CMOS circuit blocks are necessary. In regard to the early development of SiGe BiCMOS, scaling of CMOS gate length therefore rapidly progressed down to 100 nm.[27] This is because gate length of cutting-edge CMOS at that time had already been scaled to less than 50 nm. However, it still remains typically at around 0.13 μm,[41] even though scaling of CMOS is proceeding. The reason for this situation is as follows. In regard to SiGe BiCMOS technology, the most attractive feature is high-frequency performance of SiGe HBTs. Although device structures and characteristics of SiGe HBTs and CMOS are designed not to affect each other, it is difficult to integrate ultra-high-frequency SiGe HBTs with scaled CMOS when improvement of SiGe HBT performance is taken into account.

To promote LSI functionality further, it has even been proposed to combine laterally diffused MOS (LDMOS) and flash memory with SiGe BiCMOS on the same chip.[42] This combination makes it possible to form a mixed-signal system-on-a-chip (SoC) by adding high-voltage circuits for power output stages and non-volatile memories. In this case, the LDMOS module and the flash memory module are combined. As mentioned previously, to achieve a reasonable cost performance, low process complexity (such as a marginal number of masks and a minimal number of process steps) is desirable. These advances in functionality of SiGe BiCMOS stimulate modifications of the above-mentioned typical process schemes, and some mixed technologies have been developed from the viewpoint of minimizing complexity and keeping modularity.

18.4.3 Device characteristics

As described in the previous section, the fabrication process of SiGe BiCMOS is designed to maintain performance of both SiGe HBTs and CMOS. Their device characteristics are therefore almost the same as those of their base transistors. (Especially, CMOS characteristics are completely unchanged because many circuitry properties and device model parameters are used as is.) The device characteristics of SiGe HBTs and CMOS in SiGe BiCMOS are therefore not described here, and only high-quality passive elements are focused on. Such passive elements play a key role in the integration of high-speed digital functions and high-frequency RF analog functions, which is one of the most important features in applications of SiGe BiCMOS technology.

To integrate passive elements, especially high quality-factor (Q) inductors and transformers, on the same chip, a modified or additional interconnect metallization is necessary. That is, unlike standard metallization, one or two top low-resistance metal layers are typically applied, as shown in Fig. 18.9, because it is necessary to obtain low-loss performance. The low-level metal layers (near the transistors) are usually kept as the base process for high-density integration of transistors. There are many kinds of metallization, such as thick metallization, a multi-metal stack, and a low-resistivity copper, for implementing low-resistance metal layers. The choice of metallization is determined from the viewpoint of balancing the performance of passive elements with process cost.

There are many kinds of passive elements, for example resistors, capacitors, varactors, inductors, transformers, and transmission lines, and many varieties within each kind. Their details are left to already published papers related to SiGe BiCMOS technology, so only brief descriptions are given here. A variety of resistors are provided to meet the needs of a particular circuit design. The resistors are classified according to three electrical parameters, namely sheet resistance, parasitic capacitance, and temperature coefficients, and the accuracies of these parameters. According to these parameters, diffusion resistors, poly-Si resistors, or metal resistors are popularly used. Two types of capacitors, MOS and MIM (metal–insulator–metal), are typically used to satisfy the required properties, namely capacitance density, breakdown voltage, and tolerance. In the case of varactors, tunability, linearity, and quality factor are key parameters. There are two main kinds of varactors: one composed of a collector–base junction and those composed of an MOS structure. To meet the requirements of high-functionality RF circuits, it is important to integrate inductors and transformers. Key parameters of inductors are quality factor and maximum operating frequency, and those of transformers are signal transmission loss and linearity. Last but not least, passive elements include low-loss transmission lines, which are designed by microstrip or coplanar waveguide schemes depending on the desired impedance and frequency. Some typical parameters of the passive elements based on SiGe BiCMOS technology (shown in Fig. 18.9) are listed in Table 18.1.

18.5 Applications in integrated circuit (IC) and large-scale integration (LSI)

In this section, some examples of ICs and LSIs for optical transmission and microwave/millimeter-wave wireless communication systems, which are implemented as applications of the above-described SiGe HBT and BiCMOS technologies, are described.

18.5.1 Optical-fiber links

To meet the demand for expanding the transmission capacity that accompanies the rapid growth of multimedia communications, it is an effective solution to develop a high-speed optical transmission system for backbone networks. In regard to such a system, IC chips must be inexpensive to receive widespread

Table 18.1 Parameters of passive elements based on SiGe BiCMOS

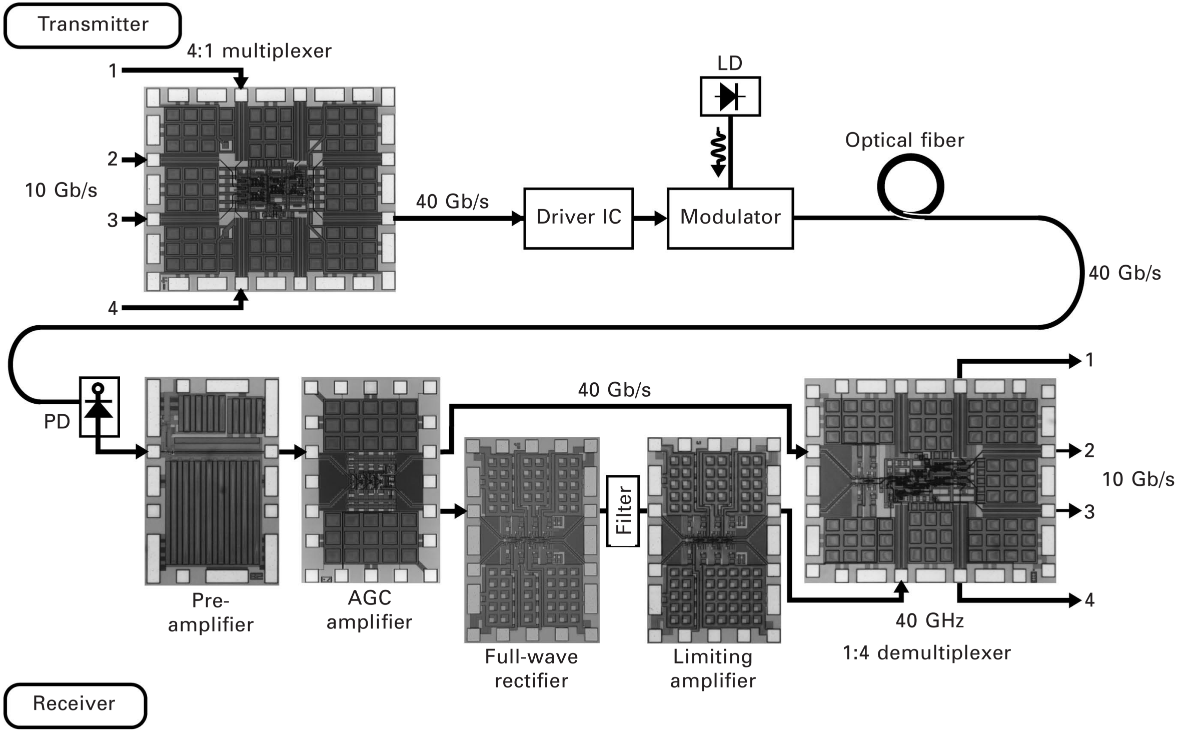
Resistor	n^+ poly Si (LR)	65	Ω/sq.
		600	ppm/°C
	p^+ poly Si (MR)	220	Ω/sq.
		330	ppm/°C
	p^+ poly Si (HR)	635	Ω/sq.
		180	ppm/°C
Capacitor	MIM	0.7	fF/μm^2
		13	Q @10 GHz
	MOS	4.7	fF/μm^2
Varactor	SiGe/Si	1.8	C_{max}/C_{min}
		45	Q @10 GHz
Inductor	Square spiral (p^- sub)	1.45	nH
		15	Q @10 GHz
	Octagonal spiral (SOI/HRS)	0.95	nH
		19	Q @10 GHz

Source: Washio, K., *IEEE Trans. Electron Dev.*, 2003, 50, 656–668, with permission.

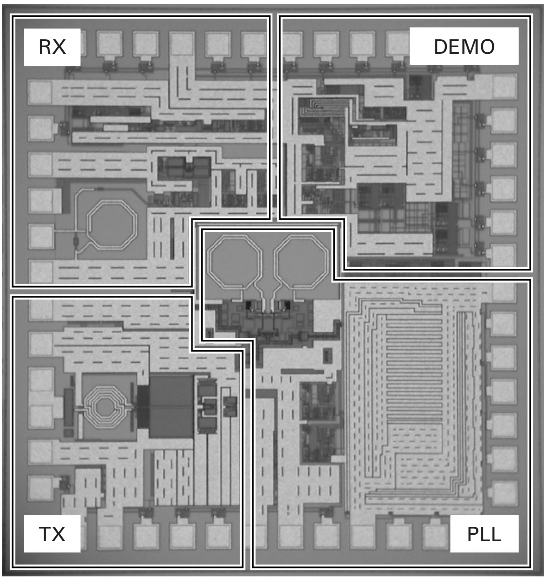
commercial approval. The SiGe HBT is therefore considered one of the most attractive candidates for such ICs. A block diagram of a transmitter and a receiver built from an IC chipset for a 40-Gb/s optical-fiber-link communication system is shown in Fig. 18.11.[1] The IC chipset includes a 4:1 multiplexer in the transmitter, and a transimpedance preamplifier, an automatic gain control (AGC) amplifier, a full-wave rectifier, a limiting amplifier, and a 1:4 demultiplexer with a decision circuit in the receiver. This configuration indicates that SiGe HBT technology with high reliability and cost-effectiveness can play an important role in optical-fiber-link systems for global communications. SiGe BiCMOS technology can provide sophisticated-function system LSIs, examples of which are a single-chip 10-Gb/s transceiver LSI[43] and a fully integrated 40-Gb/s MUX/DEMUX LSI.[44] These LSIs demonstrate that SiGe BiCMOS is a suitable technology for producing high-level integrated system LSIs in optical transmission networks.

18.5.2 Wireless communications

With the proliferation of high-performance wireless communication systems, it is necessary to ensure that all items of equipment are small and cheap. Moreover, to cut extra space and cost for assembly, a fully integrated single-chip solution is suitable. Based on SiGe BiCMOS technology, several LSIs, such as a single-chip electronic-toll-collection (ETC) IC (Fig. 18.12)[45] and a wireless LAN (IEEE 802.11a/b/g) transceiver IC,[46] have been developed. These single-chip solutions are being extended to millimeter-waveband applications, namely a 60-GHz receiver and transmitter chipset for broadband


Transmitter
4:1 multiplexer
1
2
10 Gb/s
3
4
40 Gb/s
Driver IC
LD
Modulator
Optical fiber
40 Gb/s
PD
Pre-
amplifier
AGC
amplifier
40 Gb/s
Full-wave
rectifier
Filter
Limiting
amplifier
1
2
10 Gb/s
3
4
40 GHz
1:4 demultiplexer
Receiver

18.11 Block diagram of transmitter and receiver built with the IC chipset fabricated by using the self-aligned SEG SiGe HBTs for 40-Gb/s optical-fiber-link communication systems. IC chipset includes 4:1 multiplexer in the transmitter, and transimpedance preamplifier, automatic gain control (AGC) amplifier, full-wave rectifier, limiting amplifier, and 1:4 demultiplexer with a decision circuit in the receiver. (From Ref. 1, with permission.)



18.12 Chip photomicrograph of single-chip 5.8-GHz electronic-toll-collection (ETC) transceiver IC based on SiGe BiCMOS. This fully integrated ETC chip consists of four blocks, i.e., receiver (RX), demodulator (DEMO), phase-locked loop (PLL) and transmitter (TX). (From Ref. 1, with permission.)

communications[47] and a dual-band 24-GHz/79-GHz transceiver for automotive radars.[48] Furthermore, advances in SiGe HBTs are paving the way to sub-millimeter-wave applications for high-resolution imaging, medical sensors, and wide-bandwidth communication systems. As such applications, ICs operating at over 100 GHz, namely a 165-GHz transceiver[49] and a 278-GHz oscillator,[50] have been developed. These ICs and LSIs demonstrate that SiGe HBT and BiCMOS technologies have great potential and suitability for wireless systems operating in the micro- to sub-millimeter wavebands.

18.6 Conclusion

As the key components for ICs and LSIs for multigigabit data communication systems and wide-bandwidth wireless communication systems, high-speed and high-frequency SiGe HBT and BiCMOS technologies are playing an important role. They have penetrated into a wide variety of wired and wireless communication systems. The improvements of these technologies are attributed to the progress of SiGe hetero-epitaxial growth, device structures, and fabrication processes. Their performances continue to improve, and operating frequency will reach the half THz order. Applications of SiGe HBT and BiCMOS are being extended not only to large-transmission-capacity but also to wide-bandwidth communication systems, in turn paving the way to other fields such as high-resolution radar and medical imaging. In this manner, SiGe HBT and BiCMOS technologies will therefore contribute a great deal to the welfare of society in the future.

18.7 References

1 Washio K, 'SiGe HBT and BiCMOS technologies for optical transmission and wireless communication systems', *IEEE Transactions on Electron Devices*, 2003, 50, 656–668.

2 Fischer A, Osten H-J and Richter H, 'An equilibrium model for buried SiGe strained layers', *Solid-State Electronics*, 2000, 44, 869–873.

3 Nishida A, Nakagawa K, Murakami E and Miyao M, 'Elimination of misfit dislocations in $Si_{1-x}Ge_x$/Si heterostructures by limited-area molecular-beam epitaxial growth', *Journal of Applied Physics*, 1992, 71, 5913–5917.

4 Oda K, Ohue E, Tanabe M, Shimamoto H and Washio K, '$Si_{1-x}Ge_x$ selective epitaxial growth for ultra-high-speed self-aligned HBTs', *Thin Solid Films*, 2000, 369, 358–361.

5 Kuo P, Hoyt J L, Gibbons J F, Turner J E and Lefforge D, 'Boron diffusion in Si and $Si_{1-x}Ge_x$', *Materials Research Society Symposium*, 1995, 379, 373–378.

6 Rücker H, Heinemann B, Röpke W, Kurps R, Krüger D, Lippert G and Osten H J, 'Suppressed diffusion of boron and carbon in carbon-rich silicon', *Applied Physics Letters*, 1998, 73, 1682–1684.

7 Oda K, Ohue E, Suzumura I, Hayami R, Kodama A, Shimamoto H and Washio K, 'Self-aligned selective-epitaxial-growth $Si_{1-x-y}Ge_xC_y$ HBT technology featuring 170–GHz f_{max}', *IEEE International Electron Devices Meeting*, 2001, 332–335.

8 Griglione M, Anderson T J, Haddara Y M, Law M E, Kevin S, Jones K S and van den Bogaard A, 'Diffusion of Ge in $Si_{1-x}Ge_x$/Si single quantum wells in inert and oxidizing ambients', *Journal of Applied Physics*, 2000, 88, 1366–1372.

9 Oda K, Miura M, Shimamoto H and Washio K, 'Precise control of doping profile and crystal quality improvement of SiGe HBTs using continuous epitaxial growth technology', *Thin Solid Films*, 2008, 517, 98–100.

10 Donkers J J T M, Kramer M C J C M, Van Huylenbroeck S, Choi L J, Meunier-Beillard P, Sibaja-Hernandez A, Boccardi G, van Noort W, Hurkx G A M, Vanhoucke T, Vleugels F, Winderickx G, Kunnen E, Peeters S, Baute D, De Vos B, Vandeweyer T, Loo R, Venegas R, Pijper R, Voogt F C, Decoutere S and Hijzen E A, 'A novel fully self-aligned SiGe:C HBT architecture featuring a single-step epitaxial collector-base process', *IEEE International Electron Devices Meeting*, 2007, 655–658.

11 Meister T F, Schäfer H, Franosch M, Molzer W, Aufinger K, Scheler U, Walz C, Stolz M, Boguth S and Böck J, 'SiGe bipolar technology with 74 GHz f_{max} and 11 ps gate delay', *IEEE International Electron Devices Meeting*, 1995, 739–742.

12 Washio K, Ohue E, Oda K, Tanabe M, Shimamoto H and Onai T, 'A selective–epitaxial SiGe HBT with SMI electrodes featuring 9.3–ps ECL-gate delay', *IEEE International Electron Devices Meeting*, 1997, 795–798.

13 Jagannathan B, Khater M, Pagette F, Rieh J-S, Angell D, Chen H, Florkey J, Golan F, Greenberg D R, Groves R, Jeng S J, Johnson J, Mengistu E, Schonenberg K T, Schnabel C M, Smith P, Stricker A, Ahlgren D, Freeman G, Stein K and Subbanna

S, 'Self-aligned SiGe NPN transistors with 285 GHz f_{MAX} and 207 GHz f_T in a manufacturable technology', *IEEE Electron Device Letters*, 2002, 23, 258–260.

14 Racanelli M, Schuegraf K, Kalburge A, Kar-Roy A, Shen B, Hu C, Chapek D, Howard D, Quon D, Wang F, U'ren G, Lao L, Tu H, Zheng J, Zhang J, Bell K, Yin K, Joshi P, Akhtar S, Vo S, Lee T, Shi W and Kempf P, 'Ultra high speed SiGe NPN for advanced BiCMOS technology', *IEEE International Electron Devices Meeting*, 2001, 336–339.

15 Washio K, Ohue E, Hayami R, Kodama A, Shimamoto H, Miura M, Oda K, Suzumura I, Tominari T and Hashimoto T, 'Ultra-high-speed scaled-down self-aligned SEG SiGe HBT technology', *IEEE International Electron Devices Meeting*, 2002, 767–770.

16 Meister T F, Schäfer H, Aufinger K, Stengl R, Boguth S, Schreiter R, Rest M, Knapp H, Wurzer M, Mitchell A, Böttner T and Böck J, 'SiGe bipolar technology with 3.9 ps gate delay', *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2003, 103–106.

17 Oda K, Miura M, Shimamoto H and Washio K, 'Effects of high-concentration phosphorus-doping on crystal quality and lattice strain in SiGe HBTs', *Applied Surface Science*, 2008, 254, 6017–6020.

18 Washio K, 'Size effects on DC and low-frequency-noise characteristics in epitaxially raised emitter SiGe HBTs', *IEEE Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, 2007, 162–165.

19 Washio K, Ohue E, Shimamoto H, Oda K, Hayami R, Kiyota R, Tanabe M, Kondo M, Hashimoto T and Harada T, 'A 0.2μm 180-GHz-f_{max} 6.7-ps-ECL SOI/HRS self aligned SEG SiGe HBT/CMOS technology for microwave and high-speed digital applications', *IEEE International Electron Devices Meeting*, 2000, 741–744.

20 Böck J, Meister T F, Knapp H, Zöschg D, Schäfer H, Aufinger K, Wurzer M, Boguth S, Franosch M, Stengl R, Schreiter R, Rest M and Treitinger L, 'SiGe bipolar technology for mixed digital and analogue RF applications', *IEEE International Electron Devices Meeting*, 2000, 745–748.

21 Washio K, Ohue E, Oda K, Hayami R, Tanabe M, Shimamoto H, Harada T and Kondo M, '82 GHz dynamic frequency divider in 5.5 ps ECL SiGe HBTs', *IEEE International Solid-State Circuits Conference*, 2000, 210–211.

22 Rieh J-S, Jagannathan B, Chen H, Schonenberg K T, Angell D, Chinthakindi A, Florkey J, Golan F, Greenberg D, Jeng S-J, Khater M, Pagette F, Schnabel C, Smith P, Stricker A, Vaed K, Volant R, Ahlgren D, Freeman G, Stein K and Subbanna S, 'SiGe HBTs with cut-off frequency of 350 GHz', *IEEE International Electron Devices Meeting*, 2002, 771–774.

23 http://www.itrs.net

24 Geynet B, Chevalier P, Vandelle B, Brossard F, Zerounian N, Buczko M, Gloria D, Aniel F, Dambrine G, Danneville F, Dutartre D and Chantre A, 'SiGe HBTs featuring $f_T > 400$GHz at room temperature', *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2008, 121–124.

25 Johnson E O, 'Physical limitations on frequency and power parameters of transistors', *RCA Review*, 1965, 26, 163–177.

26 Knoll D, Heinemann B, Barth R, Blum K, Borngräber J, Drews J, Ehwald K-E, Fischer G, Fox A, Grabolla T, Haak U, Höppner W, Korndörfer F, Kuck B, Marschmeyer S, Richter H, Rücker H, Schley P, Schmidt D, Scholz R, Senapati B, Tillack B, Winkler W, Wolansky D, Wolf C, Wulf H-E, Yamamoto Y and Zaumseil P, 'A modular, low-cost SiGe:C BiCMOS process featuring high-f_T and high BV_{CEO} transistors', *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2004, 241–244.

27 Washio K, 'SiGe HBT and BiCMOS technologies', *IEEE International Electron Devices Meeting*, 2003, 113–116.

28 Fox A, Heinemann B, Barth R, Bolze D, Drews J, Haak U, Knoll D, Kuck B, Kurps R, Marschmeyer S, Richter H H, Rücker H, Schley P, Schmidt D, Tillack B, Weidner G, Wipf C, Wolansky D and Yamamoto Y, 'SiGe HBT module with 2.5 ps gate delay', *IEEE International Electron Devices Meeting*, 2008, 731–734.

29 Miura M, Shimamoto H, Oda K and Washio K, 'Ultra-low-power SiGe HBT technology for wide-range microwave applications', *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2008, 129–132.

30 El-Kareh B, Balster S, Leitz W, Steinmann P, Yasuda H, Corsi M, Dawoodi K, Dirnecker C, Foglietti P, Haeusler A, Menz P, Ramin M, Scharnagl T, Schiekofer M, Schober M, Schulz U, Swanson L, Tatman D, Waitschull M, Weijtmans J W and Willis C, 'A 5V complementary-SiGe BiCMOS technology for high-speed precision analog circuits', *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2003, 211–214.

31 Heinemann B, Barth R, Bolze D, Drews J, Formanek P, Grabolla T, Haak U, Höppner W, Knoll D, Köpke K, Kuck B, Kurps R, Marschmeyer S, Richter H H, Rücker H, Schley P, Schmidt D, Winkler W, Wolansky D, Wulf H E and Yamamoto Y, 'A low-parasitic collector construction for high-speed SiGe:C HBTs', *IEEE International Electron Devices Meeting*, 2004, 251–254.

32 Knoll D, Heinemann B, Ehwald K E, Fox A, Rücker H, Barth R, Bolze D, Grabolla T, Haak U, Drews J, Kuck B, Marschmeyer S, Richter H H, Chaimanee M, Fursenko O, Schley P, Tillack B, Köpke K, Yamamoto Y, Wulf H E and Wolansky D, 'A low-cost, high-performance, high-voltage complementary BiCMOS process', *IEEE International Electron Devices Meeting*, 2006, 607–610.

33 Tominari T, Miura M, Shimamoto H, Arai M, Yoshida Y, Sato H, Aoki T, Nomami H, Wada S, Hosoe H, Washio K and Hashimoto T, 'A 10V complementary SiGe BiCMOS foundry process for high-speed and high-voltage analog applications', *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2007, 38–41.

34 Hueting R J E, Slotboom J W, Melai J, Agarwal P and Magnée P H C, 'A new trench bipolar transistor for RF applications', *IEEE Transactions on Electron Devices*, 2004, 51, 1108–1113.

35 Avenier G, Schwartzmann T, Chevalier P, Vandelle B, Rubaldo L, Dutartre D, Boissonnet L, Saguin F, Pantel R, Frégonèse S, Maneux C, Zimmer T and Chantre A, 'A self-aligned vertical HBT for thin SOI SiGeC BiCMOS', *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2005, 128–131.

36 Rücker H, Heinemann B, Barth R, Bolze D, Drews J, Fursenko O, Grabolla T, Haak U, Höppner W, Knoll D, Marschmeyer S, Mohapatra N, Richter H H, Schley P, Schmidt D, Tillack B, Weidner G, Wolansky D, Wulf H E and Yamamoto Y, 'Integration of high-performance SiGe:C HBTs with thin-film SOI CMOS', *IEEE International Electron Devices Meeting*, 2004, 239–242.

37 Hashimoto T, Nonaka Y, Saito T, Sasahara K, Tominari T, Sakai K, Tokunaga K, Fujiwara T, Wada S, Udo T, Jinbo T, Washio K and Hosoe H, 'Integration of a 0.13-μm CMOS and a high performance self-aligned SiGe HBT featuring low base resistance', *IEEE International Electron Devices Meeting*, 2002, 779–782.

38 Knoll D, Rücker H, Heinemann B, Barth R, Bauer J, Bolze D, Ehwald K E, Grabolla T, Haak U, Hunger B, Krüger D, Kurps R, Marschmeyer S, Richter H H, Schley P, Tillack B and Winkler W, 'HBT before CMOS, a new modular SiGe BiCMOS integration scheme', *IEEE International Electron Devices Meeting*, 2002,

499–502.

39 Heinemann B, Knoll D, Barth R, Bolze D, Blum K, Drews J, Ehwald K-E, Fischer G G, Köpke K, Krüger D, Kurps R, Rücker H, Schley P, Winkler W and Wulf H-E, 'Cost-effective high-performance high-voltage SiGe:C HBTs with 100 GHz f_T and $BV_{CEO} \times f_T$ products exceeding 220 VGHz', *IEEE International Electron Devices Meeting*, 2001, 348–351.

40 Knoll D, Ehwald K E, Heinemann B, Fox A, Blum K, Rücker H, Fürnhammer F, Senapati B, Barth R, Haak U, Höppner W, Drews J, Kurps R, Marschmeyer S, Richter H H, Grabolla T, Kuck B, Fursenko O, Schley P, Scholz R, Tillack B, Yamamoto Y, Köpke K, Wulf H E, Wolansky D and Winkler W, 'A flexible, low-cost, high performance SiGe:C BiCMOS process with a one-mask HBT module', *IEEE International Electron Devices Meeting*, 2002, 783–786.

41 Rücker H, Heinemann B, Winkler W, Barth R, Borngräber J, Drews J, Fischer G G, Fox A, Grabolla T, Haak U, Knoll D, Korndörfer F, Mai A, Marschmeyer S, Schley P, Schmidt D, Schmidt J, Schulz K, Tillack B, Wolansky D and Yamamoto Y, 'A 0.13µm SiGe BiCMOS technology featuring f_T/f_{max} of 240/330 GHz and gate delays below 3 ps', *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2009, 166–169.

42 Knoll D, Fox A, Ehwald K E, Heinemann B, Barth R, Fischer A, Rücker H, Schley P, Scholz R, Korndörfer F, Senapati B, Stikanov V E, Tillack B, Winkler W, Wolf C and Zaumseil P, 'A low-cost SiGe:C BiCMOS technology with embedded flash memory and complementary LDMOS module', *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2005, 132–135.

43 Ueno S, Watanabe K, Kato T, Shinohara T, Mikami K, Hashimoto T, Takai A, Washio K, Takeyari R and Harada T, 'A single-chip 10 Gb/s transceiver LSI using SiGe SOI/BiCMOS', *IEEE International Solid-State Circuits Conference*, 2001, 82–83.

44 Watanabe K, Koyama A, Harada T, Aida T, Ito A, Murata T, Yoshioka H, Sonehara M, Yamashita H, Ishikawa K, Ito M, Shiramizu N, Nakamura T, Ohhata K, Arakawa F, Kusunoki T, Chiba H, Kurihara T and Kuraishi M, 'A low-jitter 16:1 MUX and a high-sensitivity 1:16 DEMUX with integrated 39.8 to 43GHz VCO for OC-768 communication systems', *IEEE International Solid-State Circuits Conference*, 2004, 166–167.

45 Masuda T, Ohhata K, Shiramizu N, Hanazawa S, Kudoh M, Tanba Y, Takeuchi Y, Shimamoto H, Nagashima T and Washio K, 'Single-chip 5.8 GHz ETC transceiver IC with PLL and demodulation circuits using SiGe HBT/CMOS', *IEEE International Solid-State Circuits Conference*, 2002, 96–97.

46 Charlon O, Locher M, Visser H A, Duperray D, Chen J, Judson M, Landesman A L, Hritz C, Kohlschuetter U, Zhang Y, Ramesh C, Daanen A, Gao M, Haas S, Maheshwari V, Bury A, Nitsche G, Wrzyszcz A, Redman-White W, Bonakdar H, Waffaoui R E and Bracey M, 'A low-power high-performance SiGe BiCMOS 802.11a/b/g transceiver IC for cellular and bluetooth co-existence applications', *IEEE Journal of Solid-State Circuits*, 2006, 41, 1503–1512.

47 Reynolds S K, Floyd B A, Pfeiffer U R, Beukema T, Grzyb J, Haymes C, Gaucher B and Soyuer M, 'A silicon 60-GHz receiver and transmitter chipset for broadband communications', *IEEE Journal of Solid-State Circuits*, 2006, 41, 2820–2831.

48 Jain V, Tzeng F, Zhou L and Heydari P, 'A single-chip dual-band 22–29–GHz/77–81–GHz BiCMOS transceiver for automotive radars', *IEEE Journal of Solid-State Circuits*, 2009, 44, 3469–3485.

49 Laskin E, Chevalier P, Chantre A, Sautreuil B and Voinigescu S P, '165-GHz transceiver in SiGe technology', *IEEE Journal of Solid-State Circuits*, 2008, 43, 1087–1100.

50 Wanner R, Lachner R, Olbrich G R and Russer P, 'A SiGe monolithically integrated 278 GHz push–push oscillator', *IEEE International Microwave Symposium*, 2007, 333–336.

19

Silicon–germanium (SiGe)-based field effect transistors (FET) and complementary metal oxide semiconductor (CMOS) technologies

S. TAKAGI, The University of Tokyo, Japan

Abstract: This chapter reviews the recent results of work on silicon–germanium (SiGe) channel metal oxide semiconductor field effect transistors (MOSFETs), which are one of the most important electronic devices using SiGe alloys, as advanced complementary metal oxide semiconductor (CMOS) devices. After describing the importance of high mobility channels, experimental results and critical issues on SiGe p-channel and n-channel MOSFETs on bulk Si substrates are introduced. Next, recent experimental results on advanced SiGe-based MOS devices with high immunity against the short channel effects of the SiGe channels to which they are applied are described.

Key words: metal oxide semiconductor field effect transistors (MOSFETs), complementary metal oxide semiconductor (CMOS) devices, mobility, short channel effects, strain, SiGe-on-insulator (SGOI).

19.1 Introduction

19.1.1 Necessity for high mobility channels

The device scaling concept, which can lead to increases in both the switching speed and the number density of MOSFETs under reasonable power consumption, has been the main guiding principle of MOS device engineering over the last 30 years. It has been recognized, however, that under the sub-100 nm regime this conventional device scaling has confronted the difficulty that three main indices associated with MOSFET performance – on-current, power consumption and short channel effects – have a trade-off relationship with each other, owing to several physical and essential limitations directly related to the device miniaturization.

For example, power supply voltage, V_{dd}, and threshold voltage, V_{th}, are critical device parameters causing this trade-off relationship. Here, the power consumption, P_{consum}, and I_{on}, can be roughly described [1] by

$$P_{\mathrm{consum}} \approx afC_{\mathrm{load}}V_{\mathrm{dd}}^{2} + I_{0} \cdot 10^{-V_{\mathrm{th}}/S}V_{\mathrm{dd}} + I_{\mathrm{leak}}V_{\mathrm{dd}} \qquad 19.1$$

$$I_{\mathrm{on}} = N_{\mathrm{s}}(V_{\mathrm{dd}}) \cdot v\,(V_{\mathrm{dd}}) \approx C_{\mathrm{g}}(V_{\mathrm{dd}} - V_{\mathrm{th}}) \cdot v(V_{\mathrm{dd}}) \qquad 19.2$$

where $a, f, C_{\text{load}}, I_0, S, I_{\text{leak}}, N_s, C_g$ and v are a constant value, the operating frequency, the load capacitance, the drain current at V_g of V_{th}, the S factor, the total leakage current including gate and junction leakages, the induced surface carrier concentration in channels, the gate capacitance and the velocity near the source region, respectively. Here, the S factor is defined as the amount of the gate voltage necessary to change the drain current by one order of magnitude in the sub-threshold region of MOSFETs.

In order to realize low power MOSFETs, lower V_{dd}, higher V_{th}, smaller S (higher immunity for short channel effects) and lower I_{leak} are necessary. However, these requirements clearly conflict with those of higher I_{on} and also hold inconsistency themselves. According to eqn 19.2, lower V_{dd} and higher V_{th} lead to significant reduction in I_{on}. Also, thick gate oxides, T_{ox}, which are needed for reducing the direct tunneling current, and significant increase in I_{leak} under the ultra-thin gate oxide regime, decrease I_{on} and increase S, because of lower C_g. Also, the increase in substrate impurity concentration, N_{sub}, which is necessary for suppressing short channel effects in bulk MOSFETs and obtaining smaller S, causes the increase in I_{leak} due to junction tunneling current and gate-induced drain leakage current and the reduction in I_{on} due to lower mobility and resulting lower v.

As a consequence, any new device engineering to overcome these difficulties or to mitigate these stringent constraints in the trade-off relations is strongly needed to simultaneously satisfy the high performance and low power consumption under healthy device characteristics against these physical limitations. Particularly, in order to make both requirements of low power consumption and high performance compatible, the introduction of channels with high carrier mobility/velocity is of paramount importance. This is because these carrier-transport-enhanced channels can not only provide higher I_{on} due to higher v, but also reduce V_{dd} or increase T_{ox} (lower C_g) under a constant value of I_{on}, according to eqn 19.2, resulting in the reduction of the active power or the standby power.

The drive current of MOSFETs per gate width can be simply represented by $I_{\text{on}} \approx qN_s^{\text{source}} v_s$, where q is the elemental charge, N_s^{source} is the surface carrier concentration near the source edge, and v_s is the carrier velocity near source edge [2]. As the channel length becomes shorter, non-stationary transport becomes more dominant, where sufficient numbers of scattering events do not occur inside the channels. This situation has been formulated as quasi-ballistic transport by Lundstrom [2, 3]. Here,

$$I_{\text{on}} = qN_s^{\text{source}} v_{\text{inj}} \cdot \frac{1-r}{1+r} \qquad 19.3$$

where v_{inj} is the injection velocity at the top of the barrier near the source edge, and r is the back-scattering rate near the source region. Since r is directly related to μ_s, the enhancement of mobility can still be important in increasing I_{on} under the quasi-ballistic transport regime.

Furthermore, when channel length becomes much shorter, probably down to less than 10 nm in Si MOSFETs, and no carrier scattering events occur inside the channel, the carrier transport is dominated by full ballistic transport. Here, I_{on} in MOSFETs under this ballistic transport, also formulated by Natori [4, 5], is simply represented by

$$I_{\mathrm{on}} = qN_{\mathrm{s}}^{\mathrm{source}} v_{\mathrm{inj}} \tag{19.4}$$

Thus, the enhancement of v_{inj} is necessary to increase I_{on} of ballistic MOSFETs. It is known here [6] that v_{inj} in both the low N_{s} region (thermal velocity) and the high N_{s} region (Fermi velocity) increases with a decrease in m_{x} and, thus, the reduction in the effective mass along the current flow direction is a key in enhancing v_{inj} and resulting I_{on} under ballistic transport. Since low effective mass materials exhibit high carrier mobility as a general trend, low effective mass materials would be basically preferable from the viewpoint of both quasi-ballistic and full-ballistic transport regimes. As a consequence, strong attention has recently been paid to high current drive MOSFETs using novel carrier-transport-enhanced channels such as strained Si, SiGe, Ge and III–V MOSFETs. Among them, SiGe has been regarded as the channel material most easily applied to CMOS under the Si platform, because SiGe can be epitaxially grown on Si substrates and has already been introduced into the source and drain of advanced MOSFETs, as described below.

Thinking of future low power MOSFETs, on the other hand, the suppression of short channel effects is also indispensable, in addition to the reduction of junction leakage current and gate leakage current, which is achievable in high-*k* gate stacks. From this viewpoint, ultra-thin body (UTB) and multi-gate structures are also quite effective in higher immunity against short channel effects. These three-dimensional gate structure MOSFETs can be helpful in reducing the power consumption of MOSFETs, because of not only the smaller S factor, but also the lower parasitic capacitance and the lower junction leakage current. In addition, ultra-thin body channels with low impurity concentrations have been recognized as giving high immunity against MOSFET performance variation, which is one of the most critical issues in terms of integrating huge numbers of MOSFETs. This is because the random dopant fluctuation of channels with high density of impurities has been a main source of the performance variation and, thus, ultra-thin body channels with low impurity concentrations can reduce the variation by suppressing short channel effects. Therefore, we also emphasize device structures combining carrier-transport-enhanced channels with UTB or multi-gate configurations, which are strongly preferred for low power applications.

Thus, recent status and progress on MOSFETs using SiGe as the channel material, including ultra-thin body and multi-gate MOSFETs, will be reviewed in this chapter.

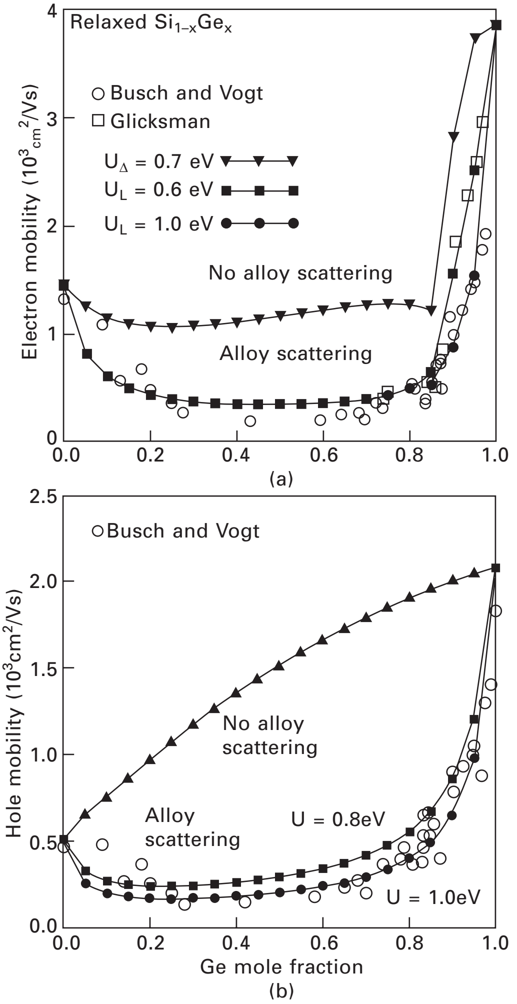
19.1.2 Roles of SiGe as channel materials and stressors in MOSFETs

As described in the previous section, the main motivation for introducing SiGe into MOSFETs is to realize high mobility (low effective mass) channel MOSFETs. Here, the roles of SiGe can be divided into two categories: one is the high mobility channel itself and the other is the stressor for creating strained-Si channels with high carrier mobility and low effective mass, utilizing the difference in the lattice constant between Si and Ge.

Channel material

As for the potential of SiGe as a high mobility channel material, high electron and high hole mobility of unstrained (relaxed) SiGe are expected for Ge contents higher than 85% under unstrained conditions, as shown in the theoretical calculations of Fig. 19.1 [7]. Both the electron and the hole mobility with alloy scattering are found to rapidly increase with an increase in the Ge fraction of SiGe in a range of the Ge fraction higher than 80%, owing to the nature of the Ge band structure, although the influence of alloy scattering on the mobility of SiGe has not yet been fully understood.

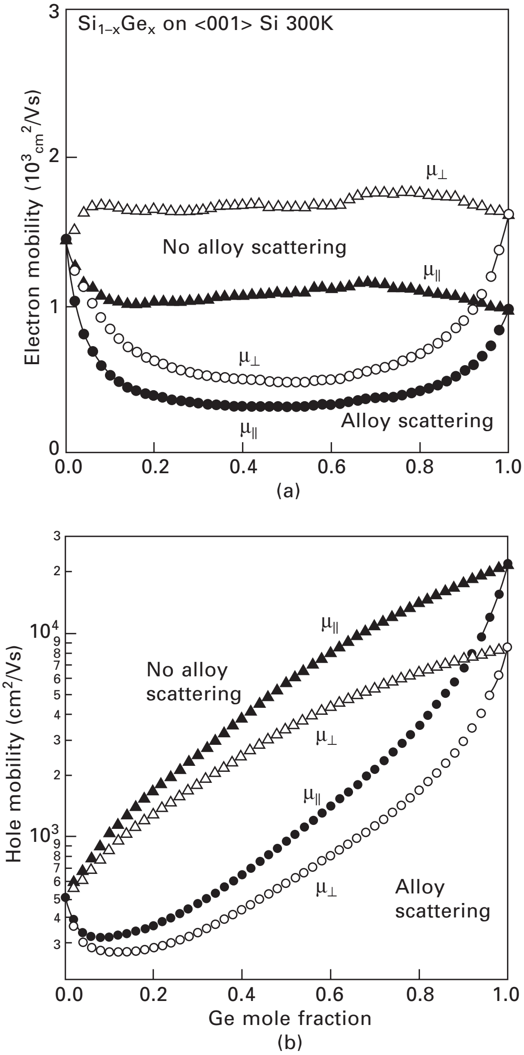
It is known, on the other hand, that the introduction of strain significantly increases the hole mobility. Figure 19.2 shows theoretical calculations of the impact of biaxial compressive strain on the bulk electron and the hole mobility [7]. It is assumed here that the SiGe layers are grown in the commensurate manner on Si substrates, meaning that the in-plane lattice constant of SiGe is the same as that of Si and, thus, higher compressive strain is applied with an increase in the Ge fraction. While the electron mobility decreases under the compressive strain, the hole mobility significantly increases up to one order of magnitude for a wide range of Ge fractions. Similarly, Fig. 19.3 shows theoretical calculations of the impact of the biaxial tensile strain of SiGe layers grown in a commensurate manner on Ge substrates [7]. In contrast to the compressive strain, the electron mobility in the Ge fraction higher than 85% is slightly enhanced, while the hole mobility is also significantly enhanced for a wide range of Ge fractions. These calculated results indicate that (1) the hole mobility is significantly enhanced in SiGe, particularly under strain, even if the Ge fraction is lower than 85%, and (2) the enhancement of the electron mobility is only observed for a Ge fraction higher than 85% even for tensile strain. As a consequence, when we introduce SiGe into MOS channels on the Si CMOS platform, where SiGe is supposed to be commensurately grown on Si substrates, p-channel MOSFET application is more straightforward for the SiGe channels.



19.1 Calculated phonon-limited 300 K (a) electron and (b) hole low field mobility in relaxed $Si_{1-x}Ge_x$ alloys. Results with and without the inclusion of alloy scattering are shown. Here, U_Δ, U_L and U are alloy scattering potential for electrons along the Δ axis (Si), electrons at the L point (Ge) and holes. Also, the open symbols were experimental data for relaxed SiGe, taken from papers by Glicksman [8] and Busch and Vogt [9].

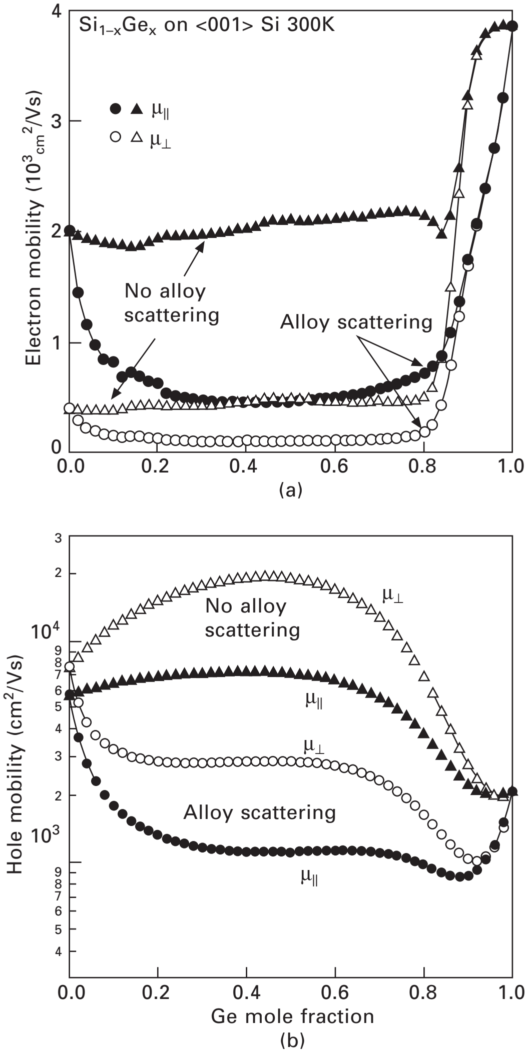
Stressor

SiGe alloys are also used in MOSFETs as stressors to introduce strain into Si and Ge channels, because the lattice constant of SiGe alloys is different from that of Si and Ge and is easily controlled by the fraction of Ge. Typical examples of the strain channels are biaxial tensile strain Si or biaxial compressive strain Ge epitaxially grown on the SiGe alloys. When any dislocations or defects



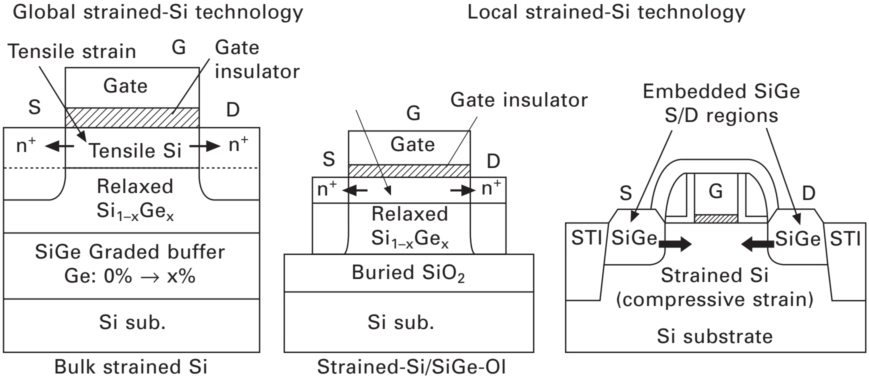
19.2 Calculated phonon-limited 300 K (a) electron and (b) hole low field mobility in $Si_{1-x}Ge_x$ alloys grown on (001) Si substrates.

are not introduced into the strained films, the amount of strain is simply in proportion to the Ge or Si fraction of the SiGe alloys. Figure 19.4 shows typical examples of SiGe stressors for creating strained-Si channels. Two approaches to introducing strained-Si channels with SiGe stressors are well known. One is called the global strain technology, where wafer-level tensile strained Si layers are formed on relaxed SiGe substrates. The structures of bulk and SOI-type global strained-Si are schematically shown in the left and center structures, respectively, in Fig. 19.4, where biaxial tensile strain can be uniformly introduced into Si MOS channels.



19.3 Calculated phonon-limited 300 K (a) electron and (b) hole low field mobility in $Si_{1-x}Ge_x$ alloys grown on (001) Ge substrates.

The other approach is called local strain technology, where materials to generate strain, such as SiGe, are locally introduced into or in the vicinity of MOSFETs during CMOS fabrication processing. While several local strain techniques are well known, one of the most effective ones is the SiGe embedded source/drain [10–13], which can introduce compressive strain into Si p-channel MOSFETs, as shown in Fig. 19.4. It is known that, from the geometrical characteristics of the embedded stressors, higher compressive stress along the channel length direction, called uniaxial strain, is introduced.


Global strained-Si technology
Local strained-Si technology
Tensile strain
G
Gate insulator
Gate
S
D
n+
Tensile Si
n+
Relaxed Si1–xGex
SiGe Graded buffer Ge: 0% → x%
Si sub.
Bulk strained Si
Gate insulator
G
Gate
S
D
n+
n+
Relaxed Si1–xGex
Buried SiO2
Si sub.
Strained-Si/SiGe-OI
Embedded SiGe S/D regions
S
G
D
STI
SiGe
SiGe
STI
Strained Si (compressive strain)
Si substrate

19.4 Schematic cross-sections of typical bulk strained-Si MOSFETs, strained-Si/SiGe-on-insulator MOSFETs and strained-Si MOSFETs with SiGe embedded source/drain.

These local strain technologies, including the SiGe-embedded source/drain, can provide a variety of advantages such as low cost, high process compatibility and easy separate optimization of n- and p-MOSFETs. Furthermore, most of these local strain technologies tend to provide uniaxial strain configurations having a strong stress component along a specific direction. The uniaxial compressive strain along the ⟨110⟩ direction can lead to a significant hole mobility enhancement in p-MOSFETs through the effective mass modulation [11, 12, 14, 15]. As a result, many studies and reports have already been devoted so far to the SiGe-embedded source/drain technologies and the MOSFETs employing SiGe as stressors [16, 17].

On the other hand, when the channel material is Ge, the type of strain becomes opposite. If planar Ge channels are epitaxially grown on relaxed SiGe substrates, the biaxial compressive strain is introduced into Ge. Also, if the embedded SiGe source/drain is formed in Ge channel MOSFETs, the uniaxial tensile strain can be introduced locally into Ge channels.

While SiGe materials have already been introduced in real LSI products as stressors and, thus, are playing an important role in advanced CMOS technologies, we will focus on the SiGe channel MOSFETs in the following sections.

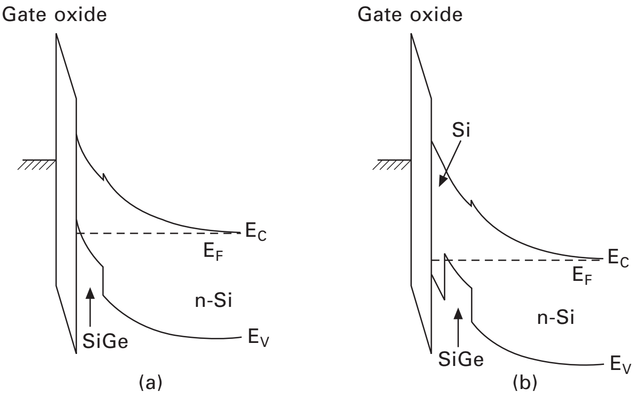
19.2 Silicon–germanium (SiGe) channel metal oxide semiconductor field effect transistors (MOSFETs)

19.2.1 SiGe p-channel MOSFETs

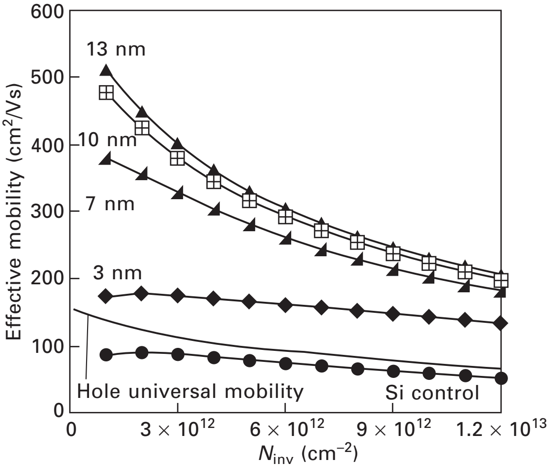
The researches on SiGe channel p-MOSFETs have a long history and there have been many reports on this device in the last few decades [18–22]. Usually, SiGe channels are epitaxially grown on Si substrates. The band diagram of surface-channel SiGe p-MOSFETs is shown in Fig. 19.5(a). When the SiGe thickness is thinner than a critical thickness, pseudomorphic SiGe with biaxial compressive stain is grown on Si. This strain is quite effective in increasing the hole mobility, as described in Section 19.1. While the increase in the Ge fraction and the resulting increase in the compressive strain effectively contribute to increasing the hole mobility, the SiGe channels need to be thinner than the critical thickness [23]. Ultrathin SiGe layers less than 10 nm thick are known to be needed for a Ge fraction higher than 0.5, even under the metastable condition.

On the other hand, it has been reported that reduction in the SiGe channel thickness leads to reduction in hole mobility [24]. Figure 19.6 shows the effective hole mobility as a function of inversion charge density, N_{inv}, for strained Si/strained $Si_{0.3}Ge_{0.7}$/relaxed $Si_{0.6}Ge_{0.4}$ structures, where p-MOSFET channels are formed in strained $Si_{0.3}Ge_{0.7}$ layers. When the SiGe thickness is less than 10 nm, the hole mobility reduction is observed at room temperature, which is attributed mainly to the increase in phonon scattering associated with the carrier confinement. Also, scattering by SiGe thickness fluctuation might also degrade the mobility in much thinner layers, though the influence of this scattering is not dominant for the SiGe thickness range examined in this study, at least down to 2.8 nm. As a result, optimization of the thickness and the Ge fraction is an important issue from the viewpoint of proper design of SiGe channels.

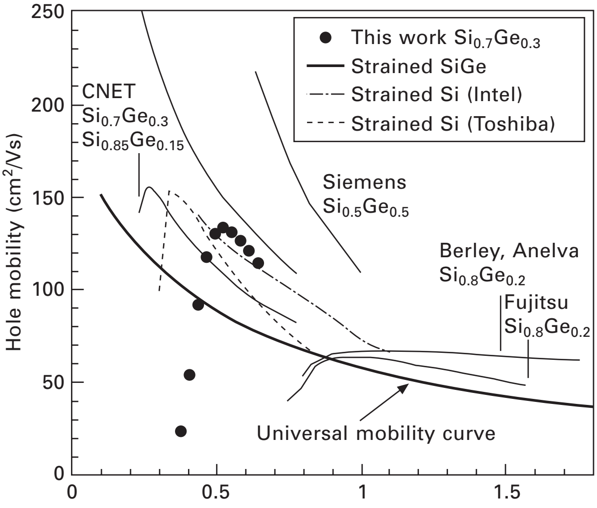
In order to mitigate this constraint between the SiGe thickness and the Ge fraction, hetero-structure MOSFETs using $Si_{1-x}Ge_x$ channels grown on relaxed $Si_{1-y}Ge_y$ virtual substrates ($x > y$) have been developed, as already shown in Fig. 19.7. Because of the relaxed $Si_{1-y}Ge_y$ buffer layers, the amount



19.5 Schematic band diagrams of SiGe MOS structures with (a) a gate oxide/SiGe direct MOS interface and (b) a gate oxide/Si/SiGe MOS interface.



19.6 Effective hole mobility as a function of inversion charge density N_{inv} for strained Si/strained $Si_{0.3}Ge_{0.7}$/relaxed $Si_{0.6}Ge_{0.4}$ structures, extracted on 50 μm × 50 μm devices using the split C–V technique. Si cap layer thickness is 2–3 nm for all devices.



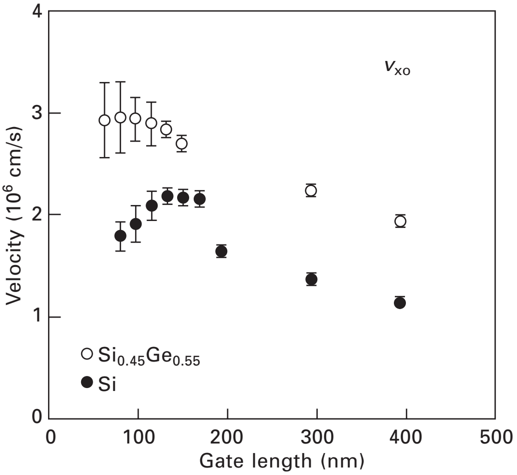
19.7 Extracted hole mobilities for p-MOSFETs containing strained $Si_{0.7}Ge_{0.3}$, and comparison with best-known data from literature (strained SiGe and strained Si). The universal mobility curve is shown as a reference. Here hole mobility data was measured on long channel devices by the split CV method. E_{eff} was extracted from CV measurements on large-capacitance structures.

of compressive strain of psuedomorphic $Si_{1-x}Ge_x$ channels is reduced and, thus, the degradation of the crystal quality due to defects and dislocations caused by the strain relaxation is expected to be suppressed. Much higher hole mobility values have been reported for strained $Si_{1-x}Ge_x$/relaxed $Si_{1-y}Ge_y$ p-MOSFETs than those for strained SiGe/Si p-MOSFETs [19, 24].

On the other hand, one of the most critical issues for SiGe MOSFETs is the poor interface properties of SiGe with gate insulators. In most cases, the interface between SiGe and gate insulators includes a large amount of interface states, and any gate stack technologies to realize the superior interface properties with SiGe have not yet been established. Therefore, thin Si capping layers are often inserted between SiGe channels and gate insulators. The band diagram of this structure is schematically shown in Fig. 19.5(b). While the existence of this Si capping layer can significantly improve the MOS interface properties, the increase in equivalent oxide thickness, resulting degradation in trans-conductance and poor immunity against short channel effects, due to the nature of buried channels, tend to be introduced simultaneously. In addition, when the gate voltage and the resulting gate oxide field are sufficiently high, the inversion channel is also formed at the capping Si surfaces and the mobility is significantly reduced down to the Si inversion layer mobility. Since these side effects become more evident in a thicker capping Si layer, the Si capping layer needs to be as thin as possible to the thickness where the surface properties are not degraded. Also, when the SiGe channel layer is very thin in this Si/SiGe/Si sandwich structure, this SiGe channel composed of the hetero-structure is sometimes called a ‘quantum well channel’.

Figure 19.7 shows several experimental data for inversion-layer hole mobility in SiGe channels on Si substrates [25]. Also, it has been reported [26] that the experimental hole mobility in $Si/Si_{0.5}Ge_{0.5}/Si$ quantum well p-MOSFETs exhibits peak values higher than 160 cm^2/Vs, while the control Si p-MOSFETs provide hole mobility of around 60 cm^2/Vs. It is found, as a result, that a hole mobility more than twice that of the universal hole mobility can be obtained for SiGe channels with high Ge fractions. The enhancement of drive current is also observed for short-channel p-MOSFETs [25, 27–29]. It has been reported [28] that an I_{on} increase of 14% and a mobility enhancement factor of 52% against control Si p-MOSFETs have been obtained in 60-nm gate length SiGe p-MOSFETs with SiN compressive stress liner, where the channel structure was composed of capping Si (0.9 nm) SiGe layers with Ge fraction of 0.225 on Si substrates.

In order to obtain high current drive in ultra-short channel MOSFETs, on the other hand, high injection velocity is needed, as described in Section 19.1.1. Figure 19.8 shows the experimental values of injection velocity evaluated for strained Si/strained $Si_{0.3}Ge_{0.7}$/relaxed $Si_{0.6}Ge_{0.4}$ p-MOSFETs [30]. It is confirmed that a higher injection velocity than in Si channels can



19.8 Comparison of the average injection velocity, v_{xo}, with extracted hole velocities. The open symbols represent strained-$Si_{0.45}Ge_{0.55}$ channel devices, while the closed symbols represent the Si control devices. The strained-SiGe and Si control devices examined here have an average L_{gate} = 150 nm and DIBL = 140 mV/V. Error bars have been added to indicated the 1 σ variation of the extracted velocities.

be obtained for strained-SiGe channels. Furthermore, as described in Section 19.2.3, uniaxial stain in SiGe channels is expected to improve the higher velocity and the current drive, because of the effective mass reduction.

In addition to the improved hole transport properties, the reduction in threshold voltage of p-MOSFETs due to the band lineup of SiGe is also supposed as one of the advantages of SiGe channel MOSFETs, particularly under metal gate/high-*k* gate stacks. One of the difficult challenges in advanced CMOS with metal gate/high-*k* gate stacks is the realization of the low threshold voltage of MOSFETs fabricated by the gate first process, which has not been easily obtained even for metal gate materials with a large work function for several reasons, such as Fermi level pinning, defect formation into high-*k* materials, dipole formation and so on. On the other hand, SiGe p-MOSFETs can inherently provide the lower threshold voltage by the valence band discontinuity between SiGe channels and Si under the same gate stack structure as that of Si p-MOSFETs, because the valence band edge of SiGe is higher than that of Si, as seen in Fig. 19.5, and thus the hole inversion layer of the SiGe channels is formed with lower absolute values of gate voltage. Therefore, when, as often reported for gate first processes, the adjustment of the threshold voltage of p-MOSFETs with metal gate/high-*k* gate stacks is a critical issue, the combination of SiGe channels with those metal gate/high-*k* gate stacks can be an effective solution [27, 31, 32]. The experimental threshold voltage reduction in SiGe p-MOSFETs has been reported to be 200–300 mV for a variety of metals, which is scalable with the gate dielectric thickness. As a result, SiGe channels allow us to provide freedom of adjustment of the threshold voltage of p-MOSFETs, depending on the Ge content and capping Si thickness.

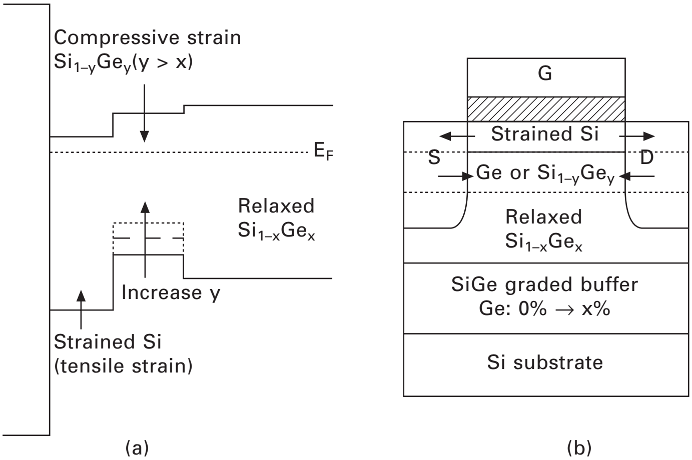
19.2.2 SiGe n-channel MOSFETs

As described in Section 19.1.1, the electron mobility of SiGe is lower than that of Ge and is also significantly degraded under compressive strain, which can commonly by introduced in SiGe channels grown on Si or SOI substrates. Also, the electron mobility of SiGe with a not too high Ge fraction is lower even than that of Si, because of the influence of alloy scattering. Thus, very few applications of SiGe to n-MOSFET channels have been examined. It has been experimentally confirmed that the effective electron mobility of SiGe is lower than that of Si [32]. Although the introduction of tensile strain to SiGe channels by using a Si embedded source/drain has been performed, the significant mobility improvement against Si n-MOSFETs has not yet been verified [34]. Therefore, pure Ge channels, which are not affected by alloy scattering, would be more promising for n-MOSFET channels than SiGe.

As a result, tensile strain Si n-MOSFETs are often combined with SiGe channel p-MOSFETs in CMOS integration. In some cases, SiGe channel layers in p-MOSFETs are used as stressors to Si channel regions. Here, there are two configurations for this integration scheme. One configuration is to grow strained-Si channel layers for n-MOSFETs directly on SiGe channels and to use this structure as CMOS, in other words, as both n-MOSFETs and p-MOSFETs. A band diagram and a schematic view of the cross-section are shown in Fig. 19.9. This CMOS structure is called 'dual hetero-channel structure CMOS' [19, 35–37]. It has been demonstrated in this structure that Ge channel p-MOSFETs with highly compressive strain exhibit a hole mobility enhancement of 10× compared with the hole mobility of Si p-MOSFETs. The other configuration is composed of the selective growth of strained-Si on SiGe layers only on the n-MOSFET regions [38–40]. In this integration scheme, the Ge fraction and the strain in the SiGe p-channels can be optimized through the additional Ge condensation, independently between n-channel and p-channel regions.

19.2.3 SiGe-on-insulator (SGOI) MOSFETs

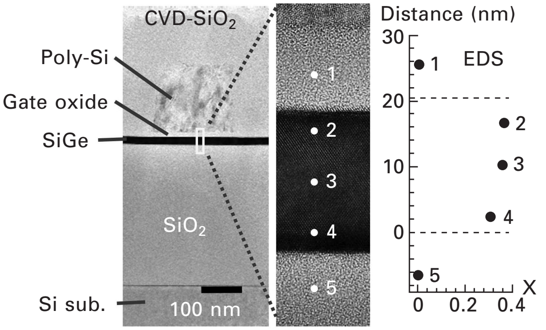
Thin body, fin or nano-wire channel structures are regarded as mandatory for future ultra-short channel MOSFETs in order to suppress the short channel effects. From this viewpoint, ultrathin-body SiGe-in-insulator (SGOI)



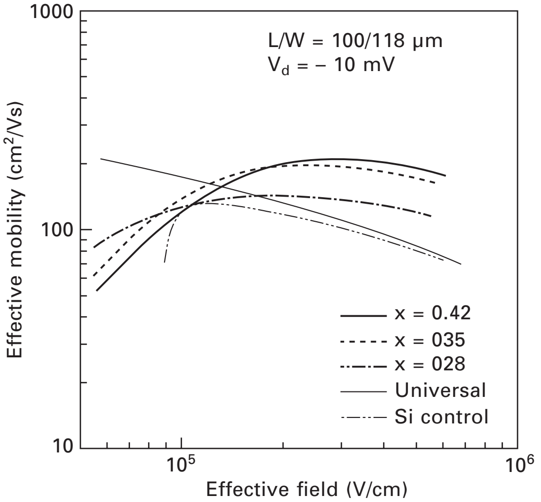
19.9 (a) Band diagram and (b) schematic cross-section of the dual hetero-channel CMOS structure.

MOSFETs or SiGe channel MOSFETs with three-dimensional gate structures such as FinFETs and nano-wires have been demonstrated by many groups. Here, the fabrication of high quality and thin SGOI layers is necessary for realizing these devices. One of the most promising techniques for preparing the SGOI layers is the Ge condensation technique [41], which is described in detail in Chapter 8. The advantages of the Ge condensation technique with respect to SGOI channel formation are listed as (1) SGOI layers can be fabricated simply by conventional processing such as thermal oxidation and SiGe epitaxial growth, (2) the Ge content and the SGOI thickness can be precisely controlled in a nanometer-order thickness range, and (3) compressive strain can be introduced to some extent during the condensation process.

Figure 19.10 shows one example of a cross-sectional TEM photograph and the Ge profile of a SGOI channel with a Ge fraction of 0.35 and a thickness of 24 nm, fabricated by the Ge condensation technique [42]. It has been confirmed here that strain relaxation does not occur up to a Ge fraction of around 0.45, meaning that this SGOI channel is compressively strained. Figure 19.11 shows the effective hole mobility of the SGOI p-MOSFETs as a function of the effective field, E_{eff}. Here, the gate insulator was formed by oxidizing the 5-nm cap Si grown on the SGOI structure down to the SiGe layer, indicating that the channel is formed at the SiO_2/SiGe interface without any Si capping layer. It is found that a mobility enhancement of 2.3 × is obtained in the high E_{eff} region for a SiGe channel with a Ge fraction of 0.43. On the other hand, the lower mobility of the SGOI p-MOSFETs in the lower E_{eff} region is attributed to the higher interface state density at the SiO_2/SiGe interface and the resulting higher Coulomb scattering.

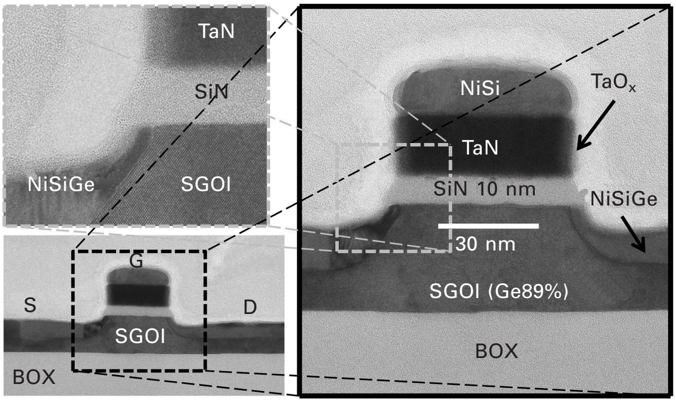


19.10 Cross-sectional TEM photographs of a SGOI channel with Ge fraction 0.35 and thickness 24 nm, fabricated by the Ge condensation technique, and the Ge fraction profile evaluated by energy dispersive X-ray spectroscopy (EDS).



19.11 Effective hole mobility of the SGOI p-MOSFETs as a function of the effective field, E_{eff}.

From the viewpoint of short channel devices, SGOI p-MOSFETs with a gate length of 60 nm and GOI p-MOSFETs with a gate length of 26 nm, fabricated by the Ge condensation technique, have been demonstrated [43]. Figure 19.12 shows TEM photographs of SGOI p-MOSFETs with a Ge fraction of 0.89 having NiSiGe metal source/drain and NiSi/TaN/SiN gate stacks. The normal operation of the 60 nm gate length SGOI p-MOSFETs and the 26 nm gate length GOI p-MOSFETs has been confirmed. On the other hand, the off-current is high, and back gate bias for suppressing the off-leakage current is needed. This high leakage current is attributable to the

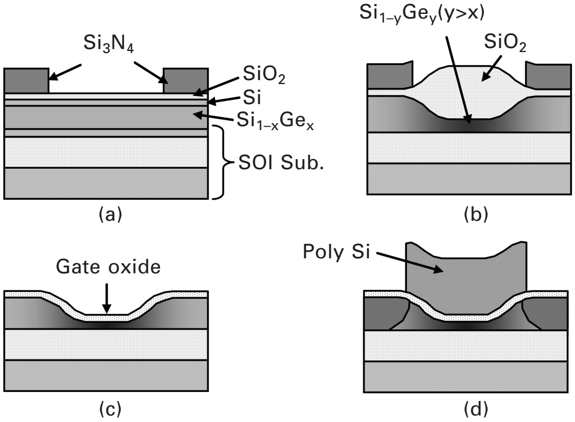


19.12 TEM micrographs of a 60 nm gate SGOI (Ge = 89%) channel metal S/D p-MOSFET.

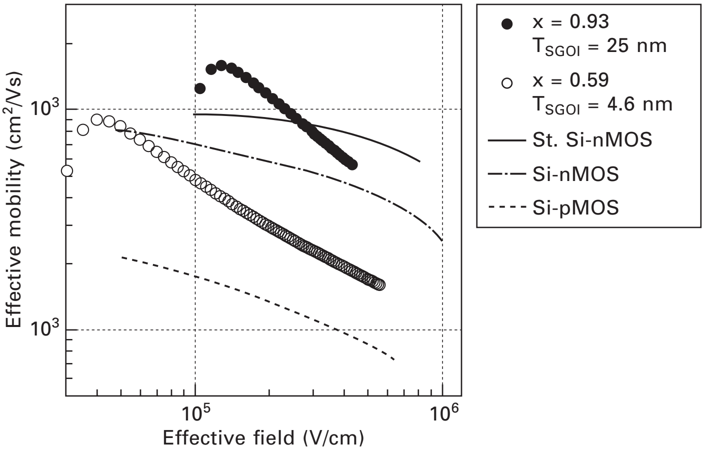
generation of a large amount of free holes, associated with the generation of any defects or dislocations created during the Ge condensation process [44]. The strain in the SGOI layers fabricated by the Ge condensation method is known to start to get relaxed around a Ge fraction higher than 0.4–0.5, which is supposed to be initiated by these many defects or dislocations [45–47]. As a result, reduction in the defect density and the resulting hole concentration, induced by the Ge condensation process, is one of the most important issues for utilizing SGOI layers fabricated by Ge condensation as MOSFET channels.

Furthermore, the concept of SGOI channel formation by the Ge condensation technique has also been applied to the selective formation of the SGOI channels [40, 48]. Figure 19.13 shows one example of the fabrication flow of compressively strained SGOI MOSFETs by the local Ge condensation technique. Here, the SGOI channels can be selectively formed in the locally oxidized regions, which is beneficial to CMOS integration using SGOI p-MOSFETs. In addition, this structure allow us to introduce the compressive strain more effectively than the planar SGOI uniformly formed on SOI substrates, because the source/drain regions with lower Ge fraction are expected to suppress the relaxation of strain in the channel regions. As described above, the strain in planar SGOI layers formed by Ge condensation is known to start to get relaxed around a Ge fraction of 0.4–0.5 [42] and to be fully relaxed near a Ge fraction of 1.0 (namely GOI) [47]. However, we have found that a substantial amount of compressive strain is still maintained near a Ge fraction of 1.0 for SGOI regions formed by local Ge condensation, though the detail of the strain relaxation behavior in the local condensation is not yet fully understood.

Figure 19.14 shows the mobility of the SGOI p-MOSFETs fabricated by



19.13 Fabrication flow of compressively strained SGOI MOSFETs by the local Ge condensation technique.



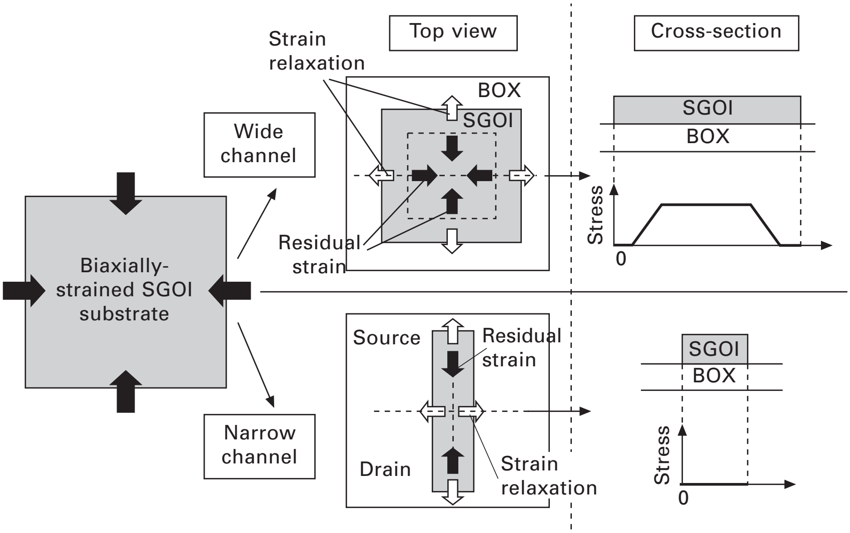
19.14 Mobility of SGOI p-MOSFETs fabricated by local condensation with Ge fractions 0.93 and 0.59 as a function of E_{eff}. For comparison, the universal electron and hole mobility in Si MOSFETs and the electron mobility of strained-Si MOSFETs with biaxial tensile strain are also shown.

local condensation as a function of E_{eff} [48]. For comparison, the universal electron and hole mobility in Si MOSFETs and the electron mobility of strained-Si MOSFETs with biaxial tensile strain are also shown. Here, the 4.6-nm-thick SGOI p-MOSFETs have a Ge fraction of 0.59 and a compressive strain of 2.5%, while the 25-nm-thick SGOI p-MOSFETs have a Ge fraction of 0.93 and a compressive strain of 1.4%, which means partial relaxation. It is found that the SGOI p-MOSFETs with Ge fractions of 0.59 and 0.89

exhibit hole mobility enhancement factors of 2 and 10, respectively, The latter mobility is higher than the universal electron mobility and almost comparable to the electron mobility in strained-Si n-MOSFETs. These results indicate that the combination of high Ge fraction and compressive strain can provide a significant increase in the hole mobility, as predicted in Fig. 19.2.

It is known, on the other hand, that for strained Si p-MOSFETs the hole mobility enhancement due to uniaxial compressive strain along the ⟨110⟩ direction is much higher than that due to biaxial strain, because of the effective mass reduction associated with the band warping [12–15]. Thus, application of uniaxial compressive strain to SiGe p-MOSFETs is also expected to provide higher mobility and velocity.

While uniaxial compressive strain can be introduced by Ge embedded source/drain, compressive stress linear or any other local strain technologies, this strain configuration can also be realized by relaxing the compressive strain only along the channel width direction for biaxial strain SiGe layers. The basic concept of realizing uniaxial strain through strain relaxation along the channel width direction is schematically shown in Fig. 19.15 [49, 50]. Fully strained SGOI substrates with biaxial compressive stress are used as starting materials and the active area (AA) is patterned by standard anisotropic etching. Since the elastic strain relaxation occurs from the edge of AA [51], transverse (lateral) stress in AA is relaxed and distributed. Thus, if AA is rectangular and its short sides are sufficiently narrow, complete and


Top view
Cross-section
Strain relaxation
BOX
SGOI
Wide channel
Residual strain
SGOI
BOX
Stress
0
Biaxially-strained SGOI substrate
Source
Residual strain
Narrow channel
Drain
Strain relaxation
SGOI
BOX
Stress
0

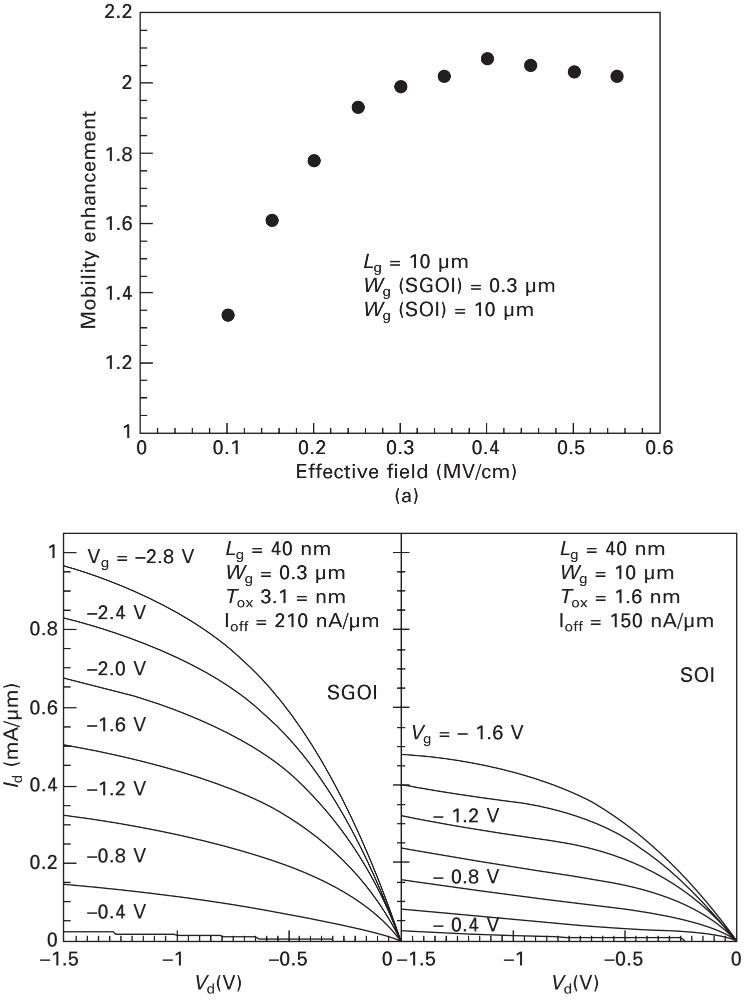
19.15 Schematic view of the basic concept of lateral strain relaxation for realizing uniaxial strain through strain relaxation along the channel width direction.

uniform lateral strain relaxation along the short side can be realized. Here, the uniaxial compressive strain along ⟨110⟩ is the best strain configuration for hole mobility enhancement not only on (100) but also on (110) [52, 53]. Therefore, when the short side of AA is taken to be the channel width direction and the channel direction is taken to be along the ⟨110⟩ direction, initial compressive stress along the channel width direction is fully relaxed, while the strain along the channel length direction along ⟨110⟩ is completely preserved. In contrast, initial biaxial compressive stress maintains inside wide and long AA. Advantages of this technique are summarized as follows: (1) large and uniform strain can be obtained, (2) no special processes are needed, (3) there is no hetero-interface that possibly causes defect formation in spite of the introduction of SiGe (Ge), and (4) the structures are highly compatible with multi-gate structures such as FinFETs and tri-gate MOSFETs.

Figure 19.16(a) shows the experimental hole mobility enhancement factor of the fabricated uniaxial SGOI p-MOSFETs with a Ge fraction of 0.2 against SOI p-MOSFETs as a function of E_{eff} [49, 50]. The mobility enhancement factor reaches 2×. This value is much higher than the hole mobility enhancement of biaxial SiGe p-MOSFETs with a low Ge fraction of 0.2 (only 10%), demonstrating the effectiveness of uniaxial compressive stress for hole mobility enhancement. The high mobility enhancement at high E_{eff} is a common character in uniaxially strained hole mobility [54]. The physical mechanism of this mobility enhancement is attributable to the effective mass reduction, similarly to that in uniaxially strained Si p-MOSFETs. Figure 19.16(b) compares the I_d–V_d characteristics of uniaxially strained SGOI and control SOI devices with L_g of 40 nm. When taking into account the difference in gate oxide thickness between SGOI and SOI, the enhancement of the current drive of uniaxially strained SGOI p-MOSFETs amounts to 1.8× against the control SOI, which is almost the same enhancement factor of the mobility as shown in Fig. 19.16(a) [50, 55]. The closeness of the enhancement factor between the low field transport (mobility) and high field transport (velocity in short channel MOSFETs) is further evidence that the effective mass reduction is the main cause of the hole transport enhancement in uniaxial SGOI p-MOSFETs.

19.2.4 SiGe FinFET, tri-gate MOSFETs

When SiGe channels are applied to ultra-short channel MOSFETs, the immunity against the short channel effect is mandatory. Thus, in addition to the ultrathin body planar MOSFETs, the application to three-dimensional multi-gate (MG) channel structures such as FinFETs, tri-gate MOSFETs and nano-wire MOSFETs is promising. Actually, the SiGe FinFETs with uniaxial compressive strain have been demonstrated [56, 57]. The device structure and the strain configuration of MG p-MOSFETs are schematically

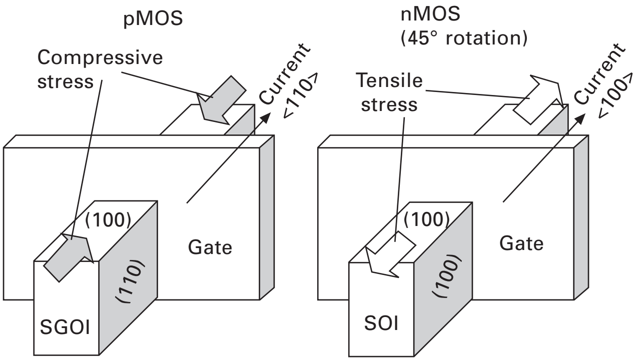


(a)

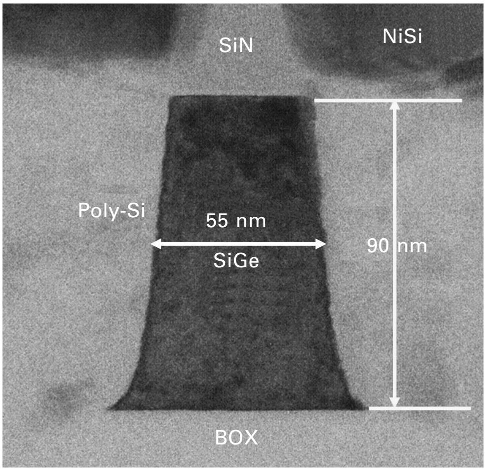
(b)

19.16 (a) Experimental hole mobility enhancement factor of the fabricated uniaxial SGOI p-MOSFETs with a Ge fraction of 0.2 against SOI p-MOSFETs as a function of E_{eff}. (b) I_d–V_d characteristics of uniaxially strained SGOI and control SOI devices with L_g of 40 nm.

shown on the left side of Fig. 19.17 [56]. Here, channel mobility boosters for pMOSFETs are known to be SiGe or Ge channels, (110) surface orientation/⟨110⟩ channel direction and uniaxial compressive strain along the ⟨110⟩ direction. In addition, lower contact and S/D sheet resistances of p^+-SiGe layers are strongly advantageous in scaled MG p-MOSFETs, where the parasitic resistance is one of the most serious problems. Thus, a main advantage of the structure shown in Fig. 19.17 is that these hole mobility


pMOS
Compressive stress
Current <110>
(100)
Gate
(110)
SGOI
nMOS (45° rotation)
Tensile stress
Current <100>
(100)
Gate
(100)
SOI

19.17 Schematic view of a three-dimensional multi-gate (MG) CMOS structure composed of a SGOI channel p-MOSFET with compressive strain and a strained-SOI channel n-MOSFET with tensile strain.


SiN
NiSi
Poly-Si
55 nm
SiGe
90 nm
BOX

19.18 TEM photograph of a fabricated SiGe FinFET.

boosters combined with the low resistance source/drain are additively integrated into the MG MOSFET structure.

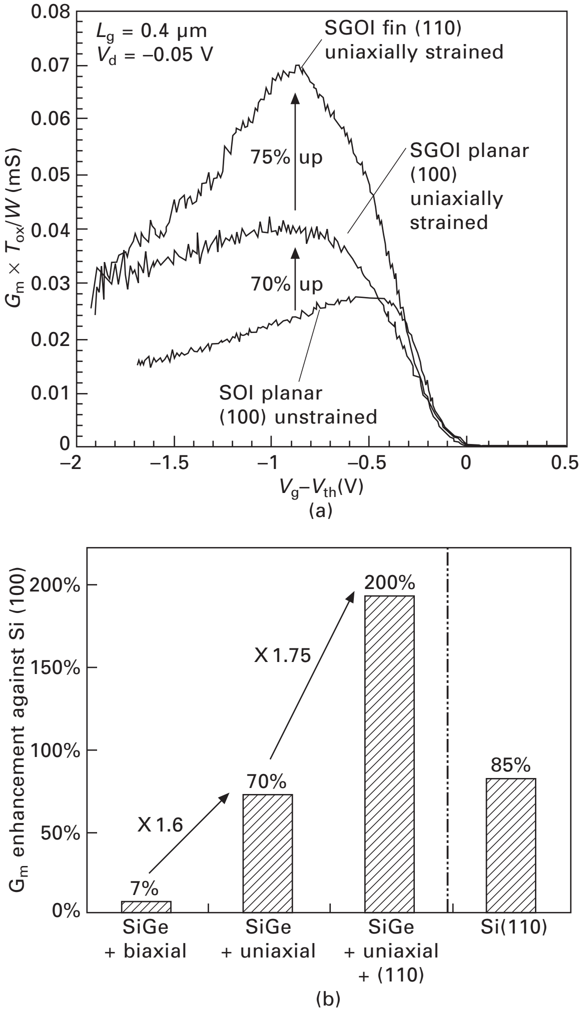
This device can be realized by applying a lateral strain relaxation process to globally strained substrates, as shown in Fig. 19.15 [56]. When fin structures are formed along the ⟨110⟩ direction on a (100) surface, the pMOS structure shown in Fig. 19.17 can be fabricated. Figure 19.18 shows a TEM photograph of the fabricated SiGe FinFETs. Under the same concept, tri-gate MOSFETs have also been demonstrated. It should be noted here that a CMOS structure with mobility enhancement due to uniaxial strain, shown in Fig. 19.17, can be realized by combining this p-MOSFET with MG n-MOSFETs along the ⟨100⟩ current flow direction with uniaxial tensile strain, which has been demonstrated by applying the same lateral strain relaxation to strained-Si-directly-on-insulator (SSOI) substrates [58, 59].

The measured trans-conductances (G_m) as a function of gate voltage among uniaxially strained SGOI FinFET, uniaxially strained planar SGOI MOSFET and planar control SOI MOSFET are compared in Fig. 19.19(a) [56]. In addition to the enhancement of G_m (70%) due to the uniaxially strained SiGe channel, another 75% enhancement due to the use of a (110) surface is observed. The decrease of the G_m enhancement in SGOI FinFET with increasing gate overdrive is attributable to the increase in surface roughness coming from non-optimized fin formation processes. Figure 19.19(b) summarizes the G_m enhancement against a Si (100) planar MOSFET resulting from each performance enhancement booster in uniaxially strained SGOI FinFETs. It is found that all the boosters employed in this device structure (SiGe, uniaxial compressive stress, (110) surface) additively contribute to the current drive, resulting in significantly high enhancement.

19.2.5 SiGe nanowire MOSFETs

Nanowire MOSFETs with much smaller channel width and surrounding gate configuration have been regarded as the ultimate CMOS structure in terms of short channel immunity. Therefore, nanowire MOSFETs using SiGe and Ge as channel materials have also been intensively studied. There are two main approaches for SiGe nanowire channel formation. One is a bottom-up approach employing CVD growth of SiGe on catalytic nanoparticles via the vapor–liquid–solid (VLS) growth mechanism [60, 61]. The other is a top-down approach using fine patterning of SiGe fin structures, pattern-dependent oxidation and so on. In this section, the examples of SiGe wire MOSFETs fabricated by the top-down approach, which is much more compatible with the Si CMOS standard processes, are briefly introduced.

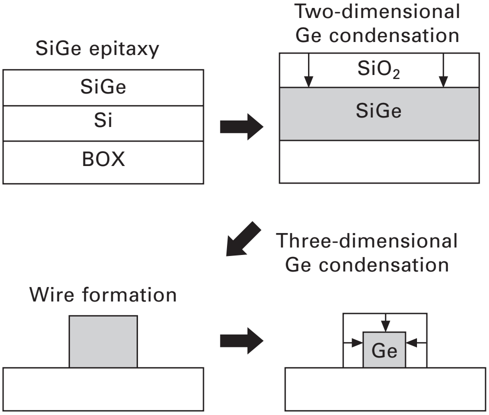
The Ge condensation process has also been used for fabricating high-Ge-fraction SiGe wire MOSFETs. A typical fabrication flow of wire-shaped SiGe MOSFETs employing three-dimensional Ge condensation is schematically shown in Fig. 19.20 [62]. First, narrow SiGe-on-insulator (SGOI) wires are patterned by standard anisotropic etching of SGOI wafers, which can be fabricated by the conventional two-dimensional Ge condensation technique. After that, the SGOI wires are oxidized (three-dimensional Ge condensation) again. In this process the width and height of SGOI wires decrease, and the Ge content inside the wires increases. The advantages of this method for fabrication of Ge nanowires are summarized as follows: (1) only the top-down and Si-compatible processes are employed, (2) the diameter and the Ge content of nanowires can be precisely controlled by the oxidation time and/or the initial wire width, and (3) defect formation accompanied by the increase of strain during the Ge condensation could be suppressed because



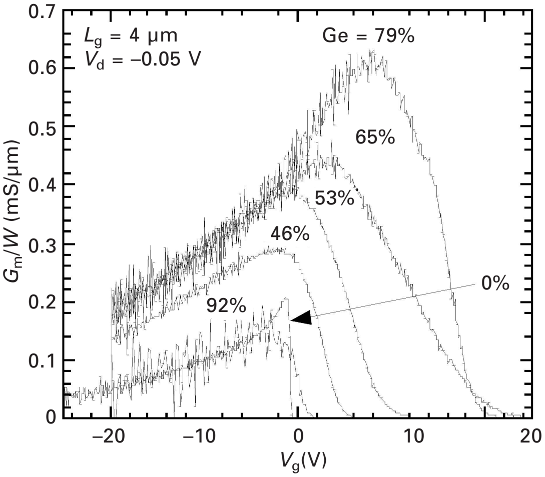
19.19 (a) Measured trans-conductance (G_m) values as a function of gate voltage among uniaxially strained SGOI FinFET, uniaxially strained planar SGOI MOSFET and planar control SOI MOSFET. (b) Amounts of G_m enhancement against a Si(100) planar MOSFET attributed to each performance enhancement booster in uniaxially strained SGOI FinFETs, SiGe channels, uniaxial compressive strain and fin structures with (110) surfaces. For comparison, the G_m enhancement factor of a Si(110) planar MOSFET is also shown.

edge-induced elastic strain relaxation is expected to take place in these small mesa structures [51].

Figure 19.21 compares the G_m–V_g characteristics of the fabricated SiGe wire p-MOSFETs with a varied Ge fraction, which is controlled by the oxidation


SiGe epitaxy
SiGe
Si
BOX
Two-dimensional Ge condensation
SiO2
SiGe
Three-dimensional Ge condensation
Wire formation
Ge

19.20 Schematic view of a typical fabrication flow of wire-shaped SiGe MOSFETs employing three-dimensional Ge condensation.

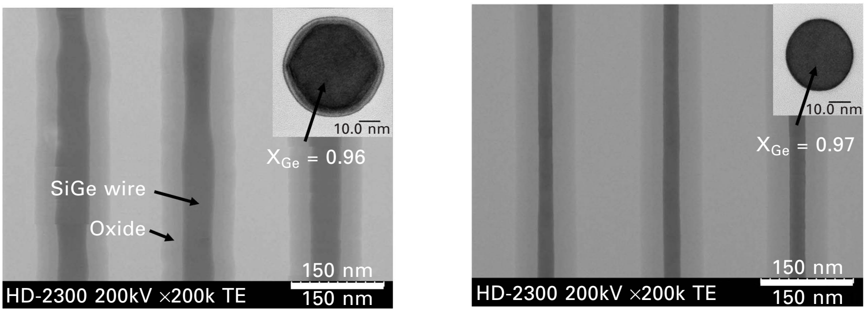


19.21 Comparison of the G_m–V_g characteristics of the fabricated SiGe wire pMOSFETs with the varied Ge fraction, which is controlled by the oxidation time.

time. The wire diameter ranges from 520 nm to 20 nm, depending on the Ge fraction. G_m is normalized by the channel width, directly evaluated by TEM pictures. It is found that G_m increases with increasing Ge fraction, except for a Ge fraction of 0.92, and that the G_m enhancement factor against the control Si device amounts to 3 at a Ge fraction of 0.79. These results show the effectiveness of three-dimensional Ge condensation for hole mobility enhancement. On the other hand, the decrease in G_m at a Ge fraction of 0.92 is attributable to the poor crystal quality due to non-optimized oxidation processes.

SiGe nanowire p-MOSFETs with a much smaller diameter of 13 nm have also been reported [63]. Here, SGOI fin structures with the 40 nm channel width are oxidized through the Ge condensation process into nanowire SGOI structures with a Ge fraction of 0.7, while the source/drain regions maintained a lower Ge fraction of 0.3 because of the wider areas, leading to compressive strain in the channel. As a result, a 4.5× enhancement factor of G_m has been obtained for SGOI nanowire MOSFETs against planar SGOI ones.

On the other hand, one of the critical issues on nanowire MOSFETs fabricated by oxidation of narrow SiGe fin structures is the control of nanowire channel diameters, which could be strongly affected by the fluctuation of the fin width due to line edge roughness in the fin structure patterning before oxidation. A trial to improve the line edge roughness and resulting fluctuation of SiGe nanowire diameters by using high temperature hydrogen etching has been reported [64, 65]. In the proposed processes, SiGe mesa structures with a hard mask are first defined by conventional lithography and RIE processes. After that, the mesas are annealed at a temperature higher than 950°C in atmospheric H_2 gas. This is an anisotropic gas etching process and result in (110) facets on their sidewalls. Successive oxidation process further trims the wires and increases the Ge fraction and strain of the SiGe wires by the Ge condensation mechanism. The top-view scanning TEM images in Fig. 19.22 show the improvement of line edge roughness of the SiGe wires after oxidation. Here, the Ge fraction is increased from 25% to 96 or 97% by the three-dimensional Ge condensation mechanism. When hydrogen etching is not used, the initial fluctuation is almost transferred to the SGOI wires after oxidation. However, hydrogen etching can realize the smooth sidewall after oxidation, demonstrating that the H_2 etching before the Gecondensation process is effective in obtaining smooth SGOI wires.



19.22 Plan-view and cross-sectional TEM images of SiGe–NWs after thermal oxidation: (a) without and (b) with the H_2 thermal etching. Here, EDX measurement revealed $x = 0.96$ and 0.97 for the SiGe wires of (a) and (b), respectively.

Hole mobility enhancement factors of 3.2× and 1.6× have been obtained for SGOI nanowire p-MOSFETs with a Ge fraction of 0.48 and a width of 35 nm against planar Si p-MOSFETs, and for SOI nanowire p-MOSFETs with a width of 28 nm, respectively.

19.3 Conclusion

The recent applications of SiGe materials to MOSFETs have been reviewed. SiGe have been mainly used for both channel materials and stressors to channels, with the latter commonly employed in the CMOS platform since the 90 nm CMOS logic technology node. In this chapter, MOSFETs using SiGe channels have been mainly reported. While the aim of the introduction of SiGe into MOS channels would be basically the same as that of Ge, SiGe channel MOSFETs could be superior to Ge ones in some cases from the viewpoints of compatibility with the standard Si CMOS platform, leakage current, strain introduction and material quality, which are still strongly dependent on the channel formation processes. Thus, device integration can be a key issue for realizing SiGe MOSFETs. As a result, SiGe channels can be a viable candidate for future high speed MOSFETs with enhanced carrier transport properties utilizing ballistic transport.

19.4 References

[1] T. Sakurai, *ISSCC Digest of Tech. Papers*, San Francisco, pp. 26–29, 2003.
[2] M. Lundstrom and J. Guo, *Nanoscale Transistors*, Springer, 2006.
[3] M. Lundstrom, *IEEE Electron Device Lett.*, vol. 22, pp. 293–295, 2001.
[4] K. Natori, *J. Appl. Phys.*, vol. 76, pp. 4879–4890, October 1994.
[5] K. Natori, *IEICE Trans. Electron.*, vol. E84-C, pp. 1029–1036, 2001.
[6] S. Takagi, *Proc. VLSI Symp.*, pp. 115–116, 2003.
[7] M. V. Fischetti and S. E. Laux, *J. Appl. Phys.*, vol. 80, pp. 2234–2252, 1996.
[8] M. Glicksman, *Phys. Rev.*, vol. 111, pp. 125–128, 1958.
[9] G. Busch and O. Vogt, *Helv. Phys. Acta.*, vol. 33, pp. 437–458, 1960.
[10] T. Ghani, M. Armstrong, C. Auth, M. Bost, P. Charvat, G. Glass, T. Hoffmann, K. Johnson, C. Kenyon, J. Klaus, B. McIntyre, K. Mistry, A. Murthy, J. Sandford, M. Silberstein, S. Sivakumar, P. Smith, K. Zawadzki, S. Thompson and M. Bohr, *Tech. Dig. IEDM*, pp. 978–981, 2003.
[11] S. E. Thompson, G. Sun, K. Wu, J. Lim and T. Nishida, *Tech. Dig. IEDM*, pp. 221–224, 2004.
[12] S. E. Thompson, M. Armstrong, C. Auth, S. Cea, R. Chau, G. Glass, T. Hoffman, J. Klaus, Z. Ma, B. Mcintyre, A. Murthy, B. Obradovic, L. Shifren, S. Sivakumar, S. Tyagi, T. Ghani, K. Mistry, M. Bohr and Y. El-Mansy, *IEEE Electron Device Lett.*, vol. 25, pp. 191–193, 2004.
[13] P. R. Chidambaram, B. A. Smith, L. H. Hall, H. Bu, S. Chakravarthi, Y. Kim, A. V. Samoilov, A. T. Kim, P. J. Jones, R. B. Irwin, M. J. Kim, A. L. P. Rotondaro, C. F. Machala and D. T. Grider, '35% drive current improvement from recessed-

SiGe drain extensions on 37 nm gate length PMOS', *Proc. VLSI Symp.*, pp. 48–49, 2004.

[14] L. Shifren, X. Wang, P. Matagne, B. Obradovic, C. Auth, S. Cea, T. Ghani, J. He, T. Hoffman, R. Kotlyar, Z. Ma, K. Mistry, R. Nagisetty, R. Shaheed, M. Stettler, C. Weber and M. D. Giles, *Appl. Phys. Lett.*, vol. 85, no. 25, pp. 6188–6190, 2004.

[15] E. X. Wang, P. Matagne, L. Shifren, B. Obradovic, R. Kotlyar, S. Cea, M. Stettle and M. D. Giles, *IEEE Trans. Electron Devices*, vol. 53, pp. 1840–1851, August 2006.

[16] P. R. Chidambaram, C. Bowen, S. Chakravarthi, C. Machala and R. Wise, *IEEE Trans. Electron Devices*, vol. 53, pp. 944–964, 2006.

[17] G. Eneman, P. Verheyen, R. Rooyackers, F. Nouri, L. Washington, R. Schreutelkamp, V. Moroz, L. Smith, A. De Keersgieter, M. Jurczak and K. De Meyer, *IEEE Trans. Electron Devices*, vol. 53, pp. 1647–1656, 2006.

[18] F. Schäffler, *Semicond. Sci. Technol.*, vol. 12, p. 1515, 1997.

[19] M. L. Lee, E. A. Fitzgerald, M. T. Bulsara, M. T. Currie and A. Lochtefeld, *J. Appl. Phys.*, vol. 97, 011101, 2005.

[20] D. K. Nayak, J. C. S. Woo, J. S. Park and K. P. MacWilliams, *IEEE Electron Device Lett.*, vol. 12, pp. 154–156, 1991.

[21] T. E. Whall and E. H. C. Parker, *J. Phys. D: Appl. Phys.*, vol. 31, p. 1397, 1998.

[22] E. H. C. Parker and T. E. Whall, *Solid-State Electron.*, vol. 43, p. 1497, 1999.

[23] P. Majhi, P. Kalra, R. Harris, K. J. Choi, D. Heh, J. Oh, D. Kelly, R. Choi, B. J. Cho, S. Banerjee, W. Tsai, H. Tseng and R. Jammy, *IEEE Electron Device Lett.*, vol. 29, pp. 99–101, 2008.

[24] C. N. Chléirigh, N. D. Theodore, H. Fukuyama, S. Mure, H.-U. Ehrke, A. Domenicucci and J. L. Hoyt, *IEEE Trans. Electron Devices*, vol. 55, pp. 2687–2694, 2008.

[25] R. Loo, N. Collaert, P. Verheyen, M. Caymax, R. Delhougne and K. De Meyer, *Appl. Surf. Sci.*, vol. 224, pp. 292–296, 2004.

[26] S.-H. Lee, J. Huang, P. Majhi, P. D. Kirsch, B.-G. Min, C.-S. Park, J. Oh, W.-Y. Loh, C.-Y. Kang, B. Sassman, P. Y. Hung, S. McCoy, J. Chen, B. Wu, G. Moori, D. Heh, C. Young, G. Bersuker, H.-H. Tseng, S. K. Banerjee and R. Jammy, *Proc. VLSI Symp.*, pp. 74–75, 2009.

[27] O. Weber, F. Ducroquet, T. Ernst, F. Andrieu, J.-F. Damlencourt, J.-M. Hartmann, B. Guillaumot, A.-M. Papon, H. Dansas, L. Brévard, A. Toffoli, P. Besson, F. Martin, Y. Morand and S. Deleonibus, *Proc. VLSI Symp.*, pp. 42–43, 2004.

[28] W.-S. Liao, Y.-G. Liaw, M.-C. Tang, K.-M. Chen, S.-Y. Huang, C.-Y. Peng and C. W. Liu, *IEEE Electron Device Lett.*, vol. 29, pp. 86–88, 2008.

[29] S.-H. Lee, P. Majhi, J. Oh, B. Sassman, C. Young, A. Bowonder, W.-Y. Loh, K.-J. Choi, B.-J. Cho, H.-D. Lee, P. Kirsch, H. R. Harris, W. Tsai, S. Datta, H.-H. Tseng, S. K. Banerjee and R. Jammy, *IEEE Electron Device Lett.*, vol. 29, pp. 1017–1019, 2008.

[30] L. Gomez, Ph.D. thesis, MIT, June 2010.; see also L. Gomez, P. Hashemi and J. L. Hoyt, *IEEE Trans. Electron Devices*, vol. 56, pp. 2644–2651, 2009.

[31] H. C.-H. Wang, S. Chen Jr, M.-F. Wang, P.-Y. Tsai, C.-W. Tsai, T.-W. Wang, S. M. Ting, T.-H. Hou, P.-S. Lim, H.-J. Lin, Y. Jin, H.-J. Tao, S.-C. Chen, C. H. Diaz, M.-S. Liang and C. Hu, *Tech. Dig. IEDM*, pp. 161–164, 2004.

[32] H. R. Harris, P. Kalra, P. Majhi, M. Hussain, D. Kelly, J. Oh, D. He, C. Smith, J. Barnett, P. D. Kirsch, G. Gebara, J. Jur, D. Lichtenwalner, A. Lubow, T. P. Ma, G. Sung, S. Thompson, B. H. Lee, H.-H. Tseng and R. Jammy, *Proc. VLSI Symp.*, pp. 154–155, 2007.

[33] C. W. Leitz, M. T. Currie, M. L. Lee, Z.-Y. Cheng, D. A. Antoniadis and E. A. Fitzgerald, *J. Appl. Phys.*, vol. 92, pp. 3745–3751, 2002.

[34] G. H. Wang, E.-H. Toh, A. Du, G.-Q. Lo, G. Samudra and Y.-C. Yeo, *IEEE Electron Device Lett.*, vol. 29, pp. 77–79, 2008.

[35] M. L. Lee and E. A. Fitzgerald, *Tech. Dig. IEDM*, pp. 429–432, 2003.

[36] S. Yu, J. Jung, J. L. Hoyt and D. A. Antoniadis, *IEEE Electron Device Lett.*, vol. 25, pp. 402–404, 2004.

[37] S. H. Olsen, A. G. O'Neill, S. Chattopadhyay, L. S. Driscoll, K. S. K. Kwa, D. J. Norris, A. G. Cullis and D. J. Paul, *IEEE Trans. Electron Devices*, vol. 51, pp. 1245–1253, 2004.

[38] T. Tezuka, S. Nakaharai, Y. Moriyama, N. Hirashita, E. Toyoda, N. Sugiyama, T. Mizuno and S. Takagi, *Proc. VLSI Symp.*, pp. 80–81, 2005.

[39] L. K. Bera, M. Mukherjee-Roy, B. Abidha, A. Agarwal, W. Y. Loh, C. H. Tung, R. Kumar, A. D. Trigg, Y. L. Foo, S. Tripathy, G. Q. Lo, N. Balasubramanian and D. L. Kwong, *IEEE Electron Device Lett.*, vol. 27, pp. 350–352, 2006

[40] T. Tezuka, S. Nakaharai, Y. Moriyama, N. Hirashita, E. Toyoda, T. Numata, T. Irisawa, K. Usuda, N. Sugiyama, T. Mizuno and S. Takagi, *Semicond. Sci. Technol.*, vol. 22, pp. S93–S98, 2007.

[41] T. Tezuka, N. Sugiyama, T. Mizuno, M. Suzuki and S. Takagi, *Jpn. J. Appl. Phys.*, Part 1, vol. 40, pp. 2866–2874, 2001.

[42] T. Tezuka, N. Sugiyama, T. Mizuno and S. Takagi, *IEEE Trans. Electron Devices*, vol. 50, pp. 1328–1333, 2003.

[43] K. Ikeda, N. Taoka, Y. Yamashita, M. Harada, K. Suzuki, T. Yamamoto, N. Sugiyama and S. Takagi, *Ext. Abs. SSDM*, pp. 30–31, 2007.

[44] N. Hirashita, Y. Moriyama, S. Nakaharai, T. Irisawa, N. Sugiyama and S. Takagi, *Appl. Phys. Exp.*, vol. 1, 101401, 2008.

[45] S. W. Bedell, K. Fogel, D. K. Sadana and H. Chen, *Appl. Phys. Lett.*, vol. 85, pp. 5869–5871, 2004.

[46] T. Tezuka, Y. Moriyama, S. Nakaharai, N. Sugiyama, N. Hirashita, E. Toyoda, Y. Miyamura and S. Takagi, *Thin Solid Films*, vol. 508, pp. 251–255, 2006.

[47] S. Nakaharai, T. Tezuka, N. Hirashita, E. Toyoda, Y. Moriyama, N. Sugiyama and S. Takagi, *Semicond. Sci. Technol.*, vol. 22, pp. S103–S106, 2007.

[48] T. Tezuka, S. Nakaharai, Y. Moriyama, N. Sugiyama and S. Takagi, *IEEE Electron Device Lett*, vol. 26, pp. 243–245, 2005.

[49] T. Irisawa, T. Numata, T. Tezuka, K. Usuda, N. Hirashita, N. Sugiyama, E. Toyoda and S. Takagi, *Proc. VLSI Symp.*, pp. 78–179, 2005.

[50] T. Irisawa, T. Numata, T. Tezuka, K. Usuda, N. Hirashita, N. Sugiyama, E. Toyoda and S. Takagi, *IEEE Trans. Electron Devices*, vol. 53, pp. 2809–2815, 2006.

[51] T. Tezuka, N. Sugiyama and S. Takagi, *J. Appl. Phys.*, vol. 94, pp. 7553–7559, 2003.

[52] C. S. Smith, *Phys. Rev. B*, vol. 94, pp. 42–49, 1954.

[53] H. Irie, K. Kita, K. Kyuno and A. Toriumi, *Tech. Dig. IEDM*, pp. 225–228, 2004.

[54] T. Ghani, M. Armstrong, C. Auth, M. Bost, P. Charvat, G. Glass, T. Hoffmann, K. Johnson, C. Kenyon, J. Klaus, B. McIntyre, K. Mistry, A. Murthy, J. Sandford, M. Silberstein, S. Sivakumar, P. Smith, K. Zawadzki, S. Thompson and M. Bohr, *Tech. Dig. IEDM*, pp. 978–980, 2003.

[55] T. Tezuka, T. Irisawa, T. Numata, Y. Moriyama, N. Hirashita, E. Toyoda, K. Usuda, N. Sugiyama and S. Takagi, *Proc. VLSI Symp.*, pp. 146–147, 2006.

[56] T. Irisawa, T. Numata, T. Tezuka, K. Usuda, S. Nakaharai, N. Hirashita, N. Sugiyama, E. Toyoda and S. Takagi, *Tech. Dig. IEDM*, pp. 727–730, 2005.

[57] C. E. Smith, H. Adhikari, S.-H. Lee, B. Coss, S. Parthasarathy, C. Young, B. Sassman, M. Cruz, C. Hobbs, P. Majhi, P. D. Kirsch and R. Jammy, *Tech. Dig. IEDM*, pp. 309–312, 2009.

[58] T. Irisawa, T. Numata, T. Tezuka, K. Usuda, N. Sugiyama and S. Takagi, *IEEE Trans. Electron Devices*, vol. 55, pp. 649–654, 2008.

[59] T. Irisawa, K. Okano, T. Horiuchi, H. Itokawa, I. Mizushima, K. Usuda, T. Tezuka, N. Sugiyama and S. Takagi, *IEEE Trans. Electron Devices*, vol. 56, pp. 1651–1658, 2009.

[60] A. B. Greytak, L. J. Lauhon, M. S. Gudiksen and C. M. Lieber, *Appl. Phys. Lett.*, vol. 84, pp. 4176–4178, 2004.

[61] C.-J. Kim, J.-E. Yang, H.-S. Lee, H. M. Jang and M.-H. Joa, *Appl. Phys. Lett.*, vol. 91, 033104, 2007.

[62] T. Irisawa, T. Numata, N. Hirashita, Y. Moriyama, S. Nakaharai, T. Tezuka, N. Sugiyama and S. Takagi, *Thin Solid Films*, vol. 517, pp. 167–169, 2008.

[63] Y. Jiang, N. Singh, T. Y. Liow, W. Y. Loh, S. Balakumar, K. M. Hoe, C. H. Tung, V. Bliznetsov, S. C. Rustagi, G. Q. Lo, D. S. H. Chan and D.-L. Kwong, *IEEE Electron Device Lett.*, vol. 29, pp. 595–598, 2008.

[64] T. Tezuka, E. Toyoda, S. Nakaharai, T. Irisawa, N. Hirashita, Y. Moriyama, N. Sugiyama, N. Taoka, Y. Yamashita, O. Kiso, M. Harada, T. Yamamoto and S. Takagi, *Tech. Dig. IEDM*, pp. 887–890, 2007.

[65] T. Tezuka, E. Toyoda, T. Irisawa, N. Hirashita, Y. Moriyama, N. Sugiyama, K. Usuda and S. Takagi, *Appl. Phys. Lett.*, vol. 94, 081910, 2009.

20

High electron mobility germanium (Ge) metal oxide semiconductor field effect transistors (MOSFETs)

A. TORIUMI, The University of Tokyo, Japan

Abstract: Significant effects of the Ge gate stack formation process on MOS capacitor and MOSFET characteristics are discussed. The GeO desorption is a key to understanding and to establishing the Ge gate stack formation process. High-pressure O_2 oxidation followed by low temperature O_2 annealing, based on thermodynamic considerations in the Ge/GeO_2 system, have dramatically improved Ge/GeO_2 MIS capacitors and MOSFET characteristics. The very low interface states density observed in Ge MOS capacitors, and the very high electron mobility in Ge MOSFETs, are demonstrated.

Key words: GeO desorption, electron mobility, Ge/GeO_2 interface, high-pressure O_2, surface orientation, interface states.

20.1 Introduction

In further miniaturized field-effect transistors (FETs), electron transport should occur as ballistic or nearly ballistic transport, where the injection velocity at the virtual source is the most important parameter for estimating the driving current of the FETs. Nevertheless, the carrier mobility in the channel is still an important parameter for achieving the ballistic transport condition [1]. Thus, a Ge complementary metal-oxide-semiconductor (CMOS) is promising for beyond scaled Si-CMOS devices, because the bulk mobility values of electrons and holes in Ge are intrinsically much higher than those in Si.

A big problem for Ge CMOS at present is the performance degradation in n-MOSFETs [2]. Electron mobility in the inversion layer of Ge n-MOSFETs is significantly degraded. Therefore, it has been discussed that Ge should be used for p-MOSFET and III–V for n-MOSFET. However, considering that a big challenge against the current CMOS large-scale integration (LSI) comes from the variability or non-uniformity of the FET performance, we should pay more attention to the simplicity of the material design and processing steps in CMOS fabrication steps. Therefore, we believe that Ge should be implemented as CMOS into LSIs, and that both n-MOSFET and p-MOSFET performance should be enhanced [3].

If the electron mobility degradation might be due to the degraded Ge interface and be inherent to Ge, Ge should be used only in p-MOSFETs as discussed. To make Ge CMOS devices feasible, we therefore need to know what determines the quality of the Ge/dielectrics interface and to understand why electron mobility has so far been severely degraded in the inversion layer. In Si CMOS technology, the Si/SiO_2 interface has long been investigated and is still a strong basis even for high-*k* gate stack technology. In that sense, the Ge/GeO_2 interface should be intensively studied to make the Ge gate stack possible. Thus, in this chapter we focus on Ge n-MOSFETs with GeO_2 dielectric film.

Although source/drain engineering in addition to gate stack formation is a critical point of the MOSFET design, it has often been neglected or even ignored in the discussion of new material devices. Both contact and diffused-layer resistances in the source/drain region should degrade FET performances even if the carrier transport in the channel may be ideally ballistic. This engineering includes controls of Schottky barrier height (SBH), dopant diffusion, activation, and solid solubility, and defect control for suppressing junction leakage. Control of these issues is not easy because these properties are strongly related to very basic properties of solid state physics and chemistry inherent to Ge. Therefore, in this chapter we only focus on SBH and metal source/drain MOSFETs.

20.2 Gate stack formation

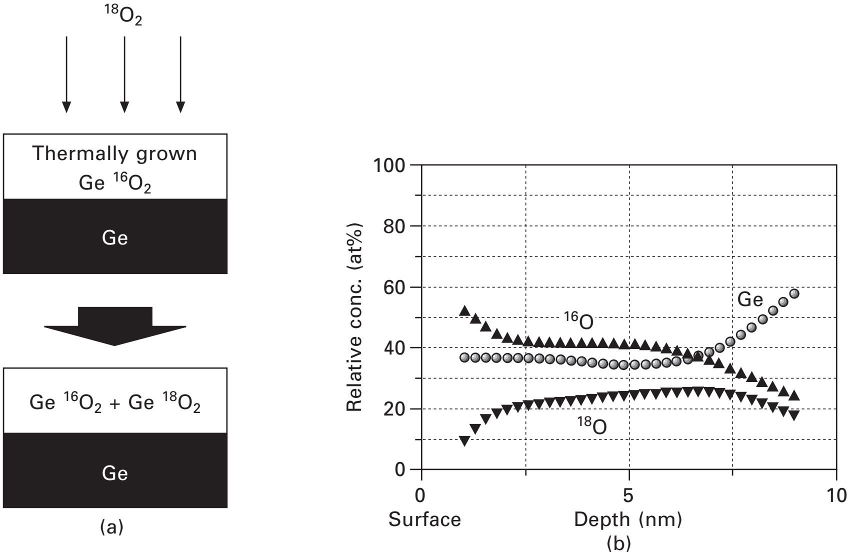
The fact that Si/SiO_2 is still the basis for Si-CMOS technology even after high-*k* has been introduced suggests the importance of the Ge/GeO_2 interface in the Ge CMOS case. Therefore, we have intensively investigated the Ge/GeO_2 interface. A big difference in the processing of Ge from that of Si is that significant GeO desorption should be taken into consideration in normal processing temperatures. This affects both the Ge/GeO_2 interface and GeO_2 bulk qualities significantly. In particular, high-temperature oxidation should be quite different from low-temperature oxidation, because in Ge oxidation GeO desorption should be seriously enhanced at higher temperatures. On the other hand, high-temperature oxidation is demanded as far as GeO_2 quality is concerned, because GeO_2 is basically in the same continuous random network structure as SiO_2 and high-temperature treatment will stabilize the GeO_2 network more.

20.2.1 Ge oxidation and GeO desorption

First, we compare the oxidation process in Ge with that in Si. Though they look the same, the oxide growth tends to saturate with oxidation time in Ge. Since we were interested in the oxidation kinetics in the Ge/GeO_2 system, we investigated the oxygen profile in GeO_2 on Ge using $^{18}O_2$ oxidation. Figure

20.1(a) shows a sample structure used in this experiment. The Ge surface was thermally oxidized by ${}^{16}O_2$, followed by ${}^{18}O_2$ at 500°C. Figure 20.1(b) shows high-resolution Rutherford backscattering spectrometry (HRBS) results where ${}^{18}O$ is not localized at the Ge/GeO$_2$ interface but is spread widely in the GeO$_2$ film (also confirmed by secondary ion mass spectrometry (SIMS), not shown). It is likely that the Deal–Grove model is roughly applicable for oxidation of Ge substrate, but this result suggests that the oxygen atoms may diffuse through oxygen atom exchange with the existing GeO$_2$ film [4]. This is quite different from Si the SiO$_2$ results [5], and is evidence that GeO$_2$ should be considered and treated differently from Si processing steps. Therefore, the oxygen kinetics in Ge oxidation should be studied in more detail for precise control of Ge/GeO$_2$ gate stacks.

The GeO desorption from GeO$_2$ was measured by thermal desorption spectroscopy (TDS). It is worth mentioning that there are typically five kinds of isotopes in natural Ge and that we need to take account of five kinds of GeO in TDS analysis. We found that GeO desorbed from GeO$_2$/Ge stacks at a relatively low temperature, but not from GeO$_2$/SiO$_2$ stacks below 700°C [6]. This indicates that GeO desorption is closely related to reactions at the Ge/GeO$_2$ interface.

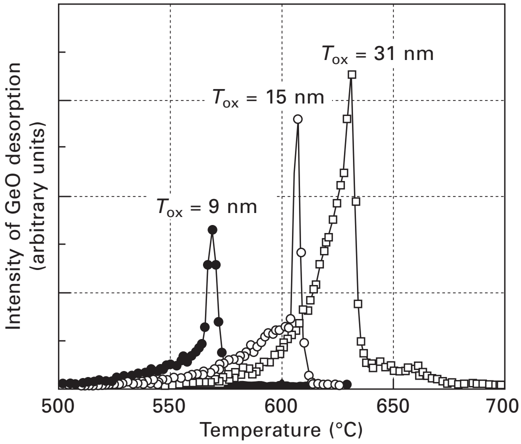


20.1 (a) Sample structure of Ge/GeO$_2$ stack for ${}^{18}O_2$ oxidation. Ge was initially oxidized by ${}^{16}O_2$, followed by ${}^{18}O_2$ oxidation. (b) Depth profiles of both ${}^{18}O$ and ${}^{16}O$ in GeO$_2$ after ${}^{18}O_2$ oxidation by high-resolution Rutherford backscattering spectrometry (HRBS). The interface profile may include artifacts due to non-uniformity or roughness, but note the relatively flat distribution of ${}^{18}O$ in GeO$_2$ film on Ge substrate.

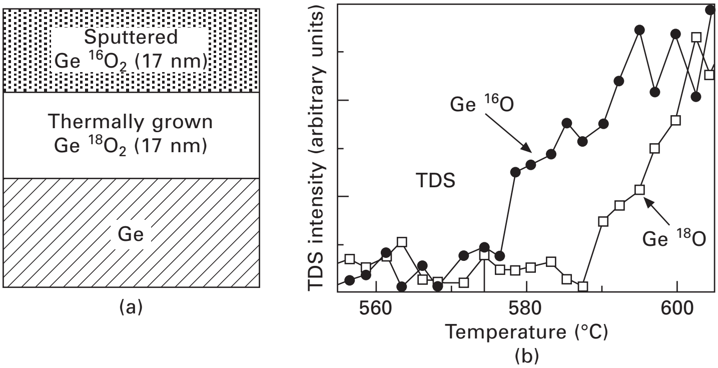
TDS measurements of thermally grown GeO_2 films on Ge (100) clarified that the GeO_2 film desorbed mainly in the form of GeO. In addition, the peak temperature depends on the film thickness, as shown in Fig. 20.2. Thicker GeO_2 film shows higher desorption temperature. This fact suggests that the GeO desorption is limited by diffusion through the GeO_2 films [7]. This is consistent with GeO desorption being observed only in the GeO_2/Ge system.

The desorption process may take place through direct GeO diffusion in the GeO_2 film from the Ge/GeO_2 interface. To clarify the kinetics of GeO desorption further, oxygen isotope (^{18}O) experiments were performed using $Ge^{16}O_2/Ge^{18}O_2$/Ge stacks [7]. TDS measurements clearly showed that GeO desorbed as $Ge^{16}O$ from the top layer in the early stage of GeO desorption with increasing temperatures, as shown in Fig. 20.3. This strongly indicates that GeO desorption is not *direct* GeO diffusion from the Ge/GeO_2 interface below 550°C in this GeO_2 thickness case, while at higher temperatures, intermixing of ^{16}O with ^{18}O may occur. Furthermore, SIMS measurements showed that the $Ge^{16}O_2/Ge^{18}O_2$ bi-layer structure was maintained even at 580°C (the TDS detected only $Ge^{16}O$ without a $Ge^{18}O$ signal at 580°C). Therefore, *direct* GeO diffusion is not likely at a realistic Ge processing temperature around 500°C.

All the experimental results thus far obtained show that GeO is formed at the Ge/GeO_2 interface. To demonstrate that the Ge wafer is directly consumed through GeO desorption, line patterns of sputtered-GeO_2 film were fabricated



20.2 Thermal desorption spectroscopy (TDS) corresponding to GeO (*m/z* = 86, 88, 89, 90, 92) from Ge/GeO_2 stacks with three kinds of initial GeO_2 thicknesses. The spectrum for thicker film shows a peak with larger intensity and at higher temperature, suggesting the diffusion process in GeO_2 film may limit GeO desorption rate. TDS measurement was performed in a UHV chamber with a lamp heating system.



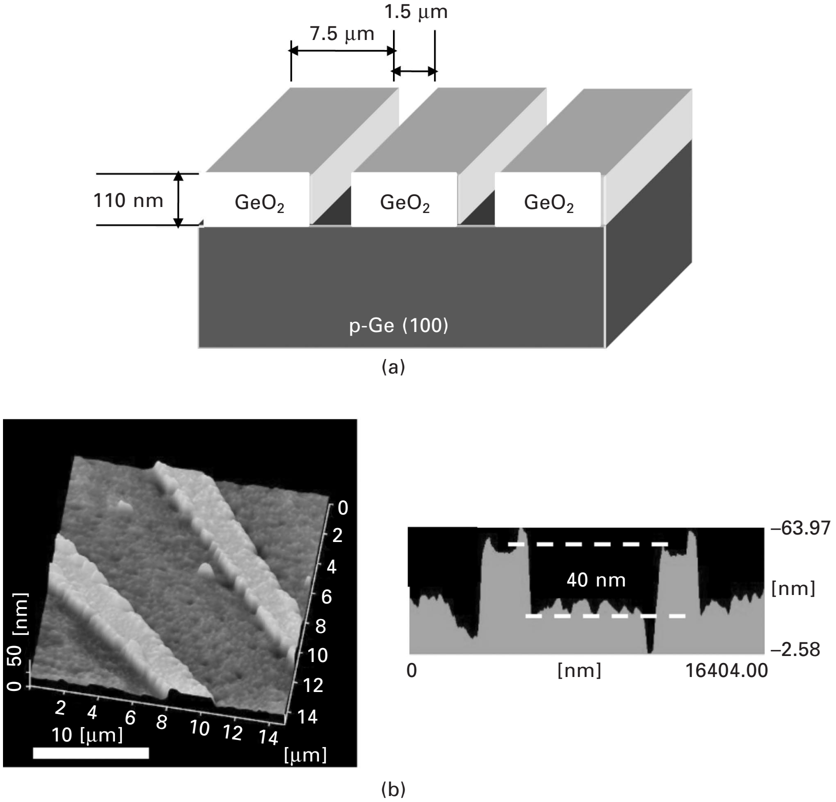
20.3 (a) Schematic of GeO_2/Ge sample with oxygen isotope (^{18}O) for TDS measurements. The bottom $Ge^{18}O_{1.3}{}^{16}O_{0.7}$ layer was thermally grown on Ge(100) in $^{18}O_2$, followed by deposition of the top $Ge^{16}O_2$ layer (by rf-sputtering). Although intentional $^{16}O_2$ was not introduced for bottom layer formation, ~30% O_2 was replaced by $^{16}O_2$. (b) TDS signals of both $^{73}Ge^{16}O$ and $^{73}Ge^{18}O$. Note that $Ge^{16}O$ desorption is dominant below 600°C. At elevated temperatures, both $Ge^{18}O$ and $Ge^{16}O$ are detected because the two layers are intermixed by oxygen diffusion.

on Ge(100) wafer as shown in Fig. 20.4(a), followed by annealing in an ultra-high vacuum (UHV). The atomic force microscopy (AFM) image in Fig. 20.4(b) reveals that the Ge wafer underneath the patterned GeO_2 films was really consumed during the UHV annealing [7, 8].

20.2.2 Thermodynamic consideration of Ge/GeO_2 system

Ge/GeO_2 interface control is thermodynamically difficult because of the high vapor pressure of GeO at the Ge/GeO_2 interface. Figure 20.5 shows the vapor pressure of GeO and GeO_2 as a function of O_2 pressure. Note that appreciable GeO vapor pressure at the GeO_2 surface is observed at 1 atm O_2 pressure at 500°C [9]. This is consistent with the fact that we detected GeO experimentally in TDS analysis as mentioned. From the viewpoint of gate stack quality, the interface property, which is determined by both GeO_2 formation and GeO desorption, should be significantly influenced by the interface reaction at Ge/GeO_2 associated with the GeO desorption. It is also quite natural that GeO_2 bulk quality is degraded in the normal oxidation process. To overcome this critical challenge, we performed Ge oxidation under high-pressure O_2 (HPO) conditions at a higher temperature to enhance the Ge oxidation as well as to thermodynamically suppress the GeO desorption.

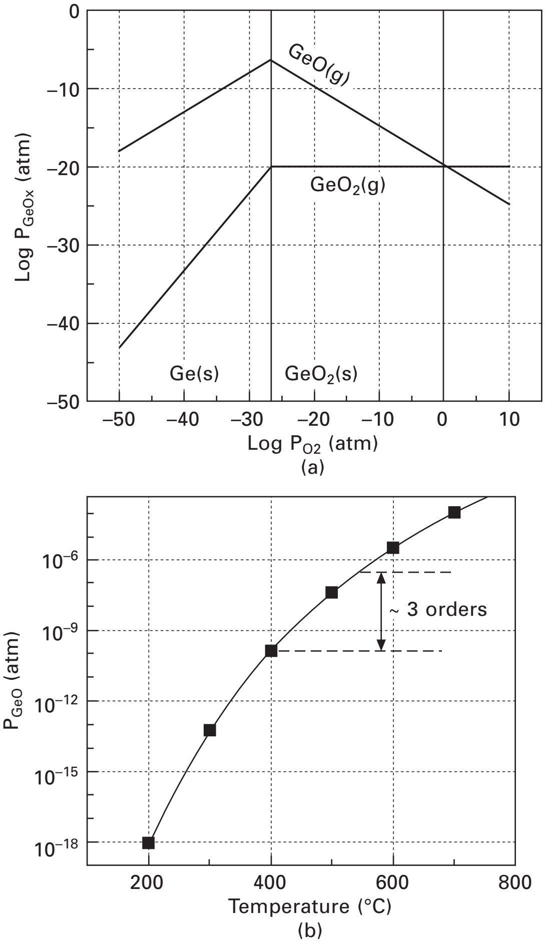
We dramatically improved MOS capacitors with GeO_2 grown by HPO [10], as shown in the room temperature (RT) C–V characteristics in Fig. 20.6. Clear depletion behavior in C–V characteristics is clearly observed



20.4 (a) Schematics of sample with line-patterned GeO_2 film on Ge, before and after annealing in UHV. (b) Atomic force microscopy (AFM) surface image and cross-sectional profile of the sample after annealing. Ge substrate was consumed by the reaction at the Ge/GeO_2 interface. This is direct evidence that GeO desorption is really triggered by the Ge/GeO_2 interface reaction.

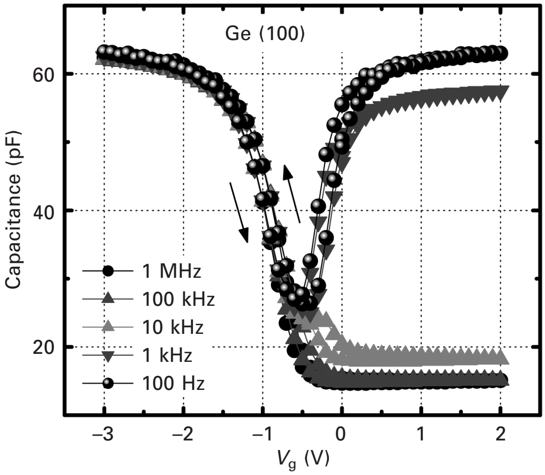
even at 100 Hz at RT with only a small hysteresis. Thus, it is concluded that HPO is a good method to enable high-temperature oxidation of Ge without deteriorating the GeO_2 bulk film quality or Ge/GeO_2 interface.

Furthermore, note that not only electrical characteristics at the interface but also physical properties of the bulk GeO_2 films grown in HPO was substantially improved, as discussed below. Figure 20.7 shows the impact of thermal treatments on GeO_2 film properties, evaluated by spectroscopic ellipsometry. In this experiment, the band edge photo-absorption of GeO_2 films was measured for differently treated GeO_2 films from the extinction coefficients (k) spectra in the spectroscopic ellipsometry measurement. An appreciable sub-gap (band edge tailing) was clearly detected after N_2 or atmospheric pressure O_2 annealing, but it disappeared after HPO [8]. This strongly suggests that sub-gap formation is correlated with oxygen deficiency

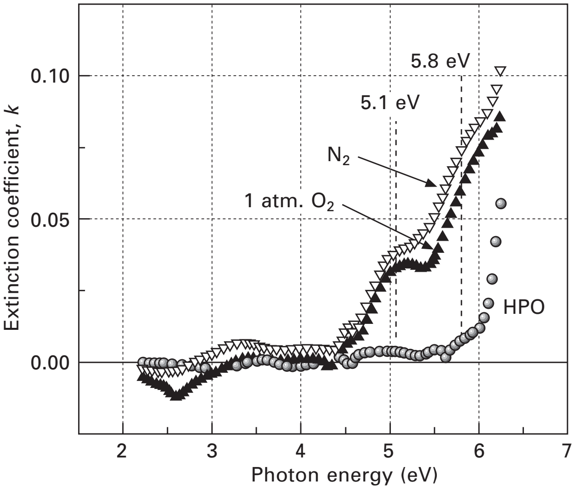


20.5 (a) Calculated GeO vapor pressure in equilibrium with GeO_2 at 550°C under various O_2 pressures. The GeO vapor pressure is highest at the Ge/GeO_2 interface. (b) Temperature dependence of the GeO vapor pressure in equilibrium with GeO_2 just at the interface. The GeO vapor pressure is very sensitive to temperature.

in the bulk GeO_2 film because the spectroscopic ellipsometry detects the bulk film properties. The sub-gap formation was observed only for GeO_2 films on Ge; it was not detected for the GeO_2 deposited on SiO_2 (data not shown). Phenomenologically, this implies that GeO desorption triggered by Ge/GeO_2 reaction at the interface may induce the oxygen deficiency in GeO_2 films. The sub-gap absorption consists of two peaks at ~5.1 and ~5.8 eV [8, 11], as indicated in Fig. 20.7. Those peaks might be correlated with defects such as neutral O vacancies or Ge^{2+}, which are reported as plausible origins of ~5-eV photo-absorption in oxygen-deficient GeO_2 bulk glass [12].



20.6 Bidirectional C–V characteristics of an Au/GeO_2/Ge(100) MOS capacitor measured at RT, where GeO_2 was grown in HPO. HPO was carried out at 550°C for 15 min. T_{ox} = 22.4 nm estimated by spectroscopic ellipsometry. 70-atm O_2 at RT before the temperature increase was used in HPO.



20.7 Extinction coefficients (*k*) estimated by spectroscopic ellipsometry of sputtered GeO_2 films on Ge(100) before and after annealing in N_2, O_2 (1 atm), and high-pressure O_2 at 600°C. Sub-gap absorption around 5–6 eV appears after annealing in both N_2 and 1 atm O_2, but does not appear at all with high pressure O_2. Sub-gap spectra can be fitted by combination of two Gaussian peaks centered at $h\nu$ = 5.1 and 5.8 eV.

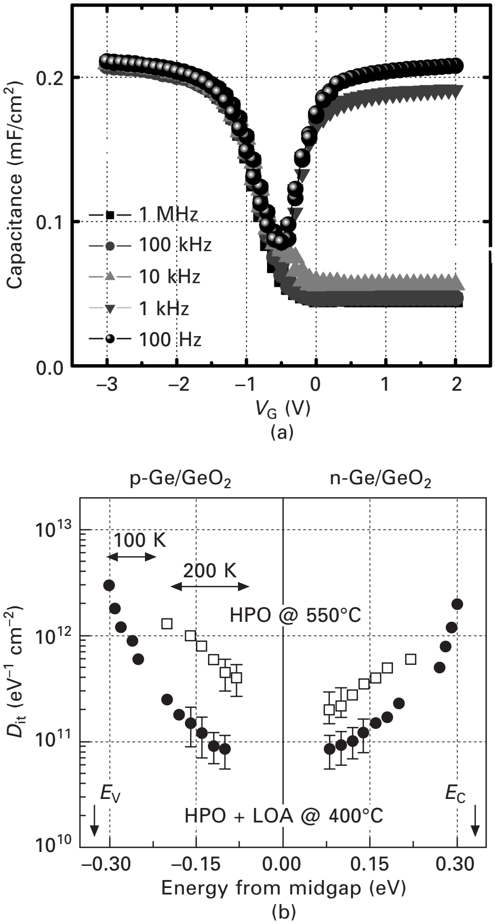
The activation energy of sub-gap formation was estimated to be ~1 eV. A similar activation energy was also extracted by the TDS results from the relationship between the starting temperature of desorption and the initial film thickness [6]. The sub-gap completely disappeared with the pO_2 (oxygen pressure) increase. The threshold pressure p_0, at which the sub-gap was annihilated, was ~40 atm at 560°C. High-temperature growth is generally required for thermal GeO_2 in terms of stabilization of the GeO_2 network formation, and a higher pO_2 is required for increasing the process temperature without sub-gap formation.

20.2.3 Au/GeO_2/Ge MOS capacitor characteristics

CV-based interface states density (D_{it}) estimation is most relevant for evaluating the GeO_2/Ge interface quality. Since the thermodynamic calculation strongly suggests that the ambient temperature affects the equilibrium pressure of GeO more strongly than pO_2 (see Fig. 20.5), a further improvement of the interface might be achieved by low-temperature O_2 annealing (LOA), because GeO vapor pressure at the interface is dramatically lowered. Then we considered the two-step oxidation process in the following. After the formation of high-quality bulk GeO_2 by HPO at 550°C, LOA was performed at 400°C in 1 atm O_2 (HPO + LOA process) [8, 13]. Figure 20.8 shows (a) C–V characteristics at RT and (b) D_{it} profiles obtained by the conductance method at 200 K, in which the mid-gap D_{it} is on the order of 10^{10} cm^{-2} eV^{-1} in HPO + LOA. This is one of the lowest D_{it} values so far reported for the GeO_2/Ge system. Note that relatively low-temperature O_2 annealing further improved the interface quality of Ge/GeO_2 without further oxidation of Ge at the interface, and that forming gas annealing was not needed to reduce D_{it} at the interface. Since the two-step process of HPO followed by LOA is based on the thermodynamic guideline, we can conclude that it is quite beneficial for Ge/GeO_2 formation. It is known that this is not effective for the Si/SiO_2 system. The microscopic origin of this advantage is now under investigation, but it may be due to more flexible interface Ge–O bonds at the Ge/GeO_2 interface compared with those at Si/SiO_2 interfaces. Thus, we can regard LOA as the 'self-passivation' process by oxygen at the Ge/GeO_2 interface instead of hydrogen passivation for the Si/SiO_2 interface [10].

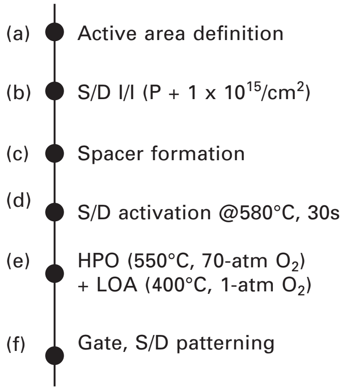
20.3 Metal oxide semiconductor field effect transistor (MOSFET) fabrication and electron inversion layer mobility

We are interested in Ge CMOS, where n-MOSFET technology is the key as mentioned in the introduction. We applied the gate stack process described in the previous section to n-MOSFET fabrication. Long channel n-MOSFETs



20.8 (a) Bi-directional C–V characteristics of Au/GeO$_2$/Ge(100) MOS capacitors at RT, where GeO$_2$ was grown in HPO followed by LOA. HPO was performed at 550°C for 15 min; LOA was performed at 400°C for 30 min in 1 atm O$_2$. (b) Energy distribution of interface states density D_{it} estimated by the low temperature conductance method at 100 and 200 K. The extracted D_{it} is lowest near the midgap and increases near the edges. The minimum D_{it} in Ge/GeO$_2$ MOS capacitors with HPO is as low as 2×10^{11} eV^{-1} cm^{-2} and is further reduced to below 10^{11} eV^{-1} cm^{-2} by adding LOA

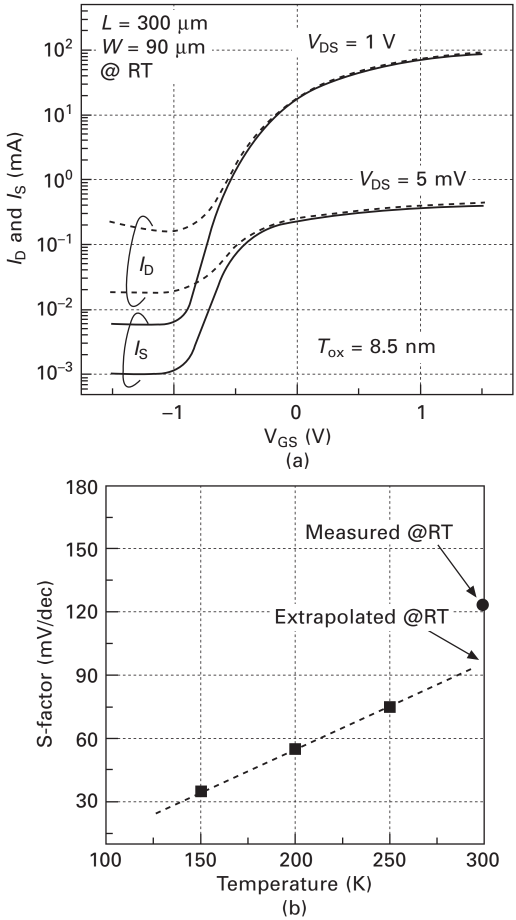
were fabricated on both Ge (100) and (111) surfaces using the two-step gate stack formation process. Figure 20.9 shows the schematic process flow of Ge n-MOSFET fabrication on both p-type Ge(100) and Ge(111) wafers (2–3 Ω cm). After the chemical cleaning of Ge substrates followed by diluted HF, 50-nm Y_2O_3 and 500-nm spin-on-glass (SOG) SiO_2 were deposited for spacer and field oxide formation, respectively. Several channel lengths


(a) Active area definition
(b) S/D I/I (P + 1 x 10^{15}/cm^2)
(c) Spacer formation
(d) S/D activation @580°C, 30s
(e) HPO (550°C, 70-atm O$_2$) + LOA (400°C, 1-atm O$_2$)
(f) Gate, S/D patterning

20.9 Schematic process flow of Ge n-MOSFETs. (a) Active area definition by SOG etching; (b) phosphorus (1 × 10^{15}/cm^2) implantation with 100 keV for S/D formation; (c) Y_2O_3 etching for spacer formation; (d) impurity activation; (e) thermal oxidation of Ge by HPO, followed by LOA; (f) gate electrode and S/D formation after Al deposition.

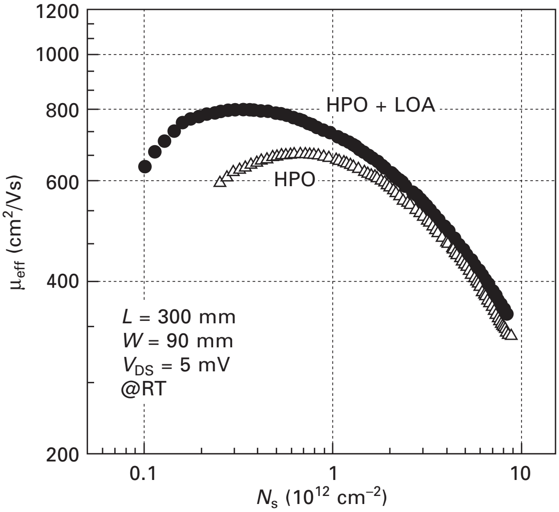
(W/L = 90 μm/100–500 μm) were defined, and phosphorus (1 × 10^{15}/cm^2 dose) was implanted with 100 keV to make the source/drain (S/D) regions. The activation of phosphorus in Ge was performed at 580°C for 30 s. Then, GeO_2 was grown at 550°C for 10 min under O_2 pressure of 70 atm at RT, followed by LOA at 400°C. Finally, Al was deposited and a gate electrode was formed. Note that Y_2O_3 spacer was used to protect the gate dielectric GeO_2 film from the wet process and air exposure because Y_2O_3 was water-resistant and very compatible with Ge [14].

We confirmed that there was very small hysteresis and frequency dispersion (data not shown) in the C–V characteristics of the fabricated n-MOSFETs. This indicates that the superiority of the Ge/GeO_2 interface properties remained even after MOSFET fabrication thanks to the interface-conscious FET process. Figure 20.10(a) shows the transfer characteristics (I_D–V_{GS} and I_S–V_{GS}) of Ge(100) n-MOSFETs at V_{DS} = 1 V and 5 mV, where 8.5 nm thick GeO_2 was grown by HPO and LOA. Since the source/drain junction formation was not optimized, quite a big leakage current at the drain side was observed. So, I_S was used for transport analysis below. The I_{on}/I_{off} current ratio was ~10^4, and the S-factor of 125 mV/dec was achieved at 300 K (V_{DS} = 1 V). This S-factor is quite a high value considering the low D_{it} extracted from the conductance method as shown in Fig. 20.8. Figure 10(b) shows the temperature dependence of the S-factor from 150 to 300 K. The estimated S-factor at 300 K is 95 mV/dec, which is reasonable and much smaller than the measured value. This inconsistency suggests that the S-factor at 300 K seems to be apparently degraded by the S/D junction leakage current, because the junction leakage is sharply decreased by lowering the measurement temperature.



20.10 (a) Measured transfer characteristics (I_S–V_{GS}) of Ge(100) n-MOSFETs for V_{DS} = 1 V and 5 mV at RT, where GeO_2 was grown by HPO and LOA. I_D (drain current) is much higher than I_S (source current) in the low V_{GS} region, because the pn junction formation process was not optimized. (b) Temperature dependence of the *S*-factor from 150 K to RT. The *S*-factor at RT estimated from temperature dependence is 95 mV/dec, which is much smaller than the measured value. This difference suggests that the *S*-factor at RT might be degraded by S/D junction leakage current.

The inversion layer electron mobility (μ_{eff}) was estimated by the split-CV method, and several gate length FETs were used to remove the parasitic resistance effect. Figure 20.11 shows the impact of LOA on μ_{eff} in addition to that of HPO as a function of inversion electron density (N_S). The peak mobility after HPO + LOA was 790 cm^2/Vs, while that after HPO was 650 cm^2/Vs without any parasitic resistance correction. The degradation of μ_{eff}

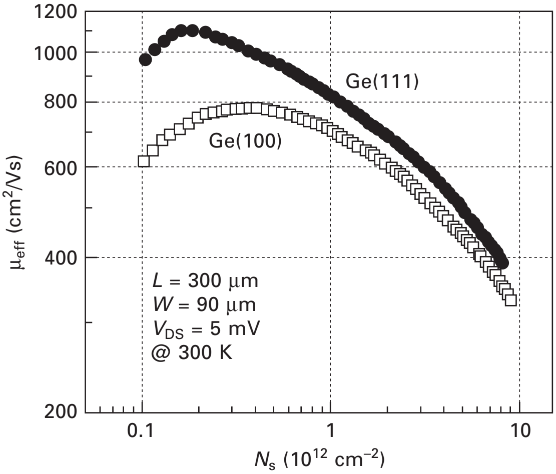


20.11 Extracted electron mobility in Ge(100) n-MOSFET as a function of inversion electron density at RT. Peak mobility after LOA is about 790 cm^2/Vs, while that just after HPO is 650 cm^2/Vs. A smaller μ_{eff} in the case of only the HPO process is explainable by Coulomb scattering due to higher values of D_{it} as shown in Fig. *20.8*. Combination of HPO and LOA is very effective for enhancing electron mobility and is almost comparable to Si universal mobility on Si (100).

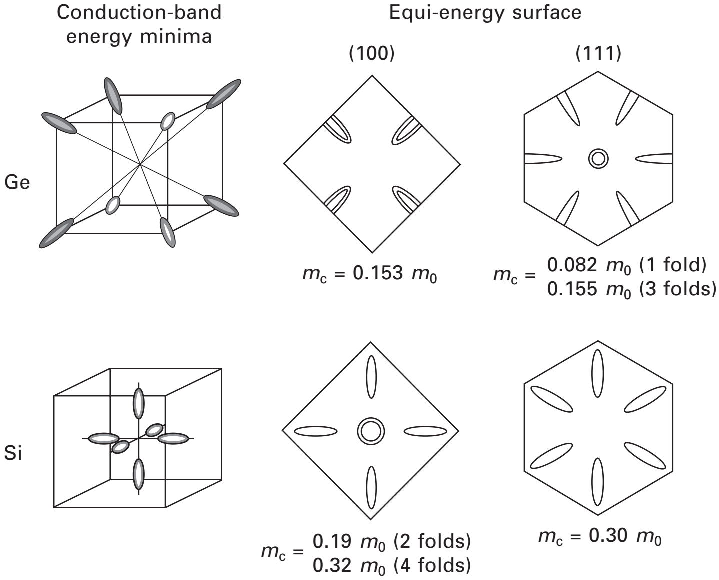
in HPO is explainable by Coulomb scattering due to higher values of D_{it} in HPO as shown in Fig. 20.8. Note that the combination of HPO with LOA is very effective for enhancing electron mobility and that μ_{eff} is comparable to the universal electron mobility in Si n-MOSFETs.

Next, the substrate orientation effect on μ_{eff} is shown on Ge(100) and Ge(111) in Fig. 20.12. The peak mobility on Ge(111) was about 1100 cm^2/Vs at RT, a 1.5-times improvement compared to that on Ge(100). This value is the highest μ_{eff} so far reported in the literature concerning Ge MOSFETs. This is quite interesting by considering the difference in the Fermi surface of Ge(100) and (111), as shown in Fig. 20.13. It suggests that the (111) surface should intrinsically be better than the (100) one in terms of the lower effective mass in the two-dimensional inversion layer. Based on this concept, the advantage of MOSFETs on the Ge(111) surface under ballistic carrier transport has also been analytically expected [15].

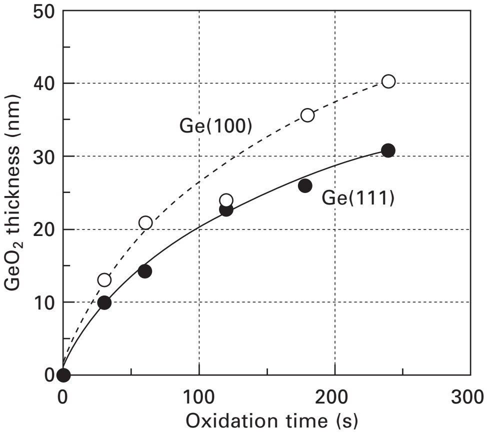
Figure 20.14(a) shows GeO_2 thickness on Ge(100) and Ge(111) grown in 1 atm O_2 at 600°C as a function of oxidation time. The oxidation rate of Ge(111) was found to be much slower than that of (100). This suggests that Ge(111) should be better in terms of GeO_2 scalability as an interfacial layer of high-*k* material. Furthermore, the substrate orientation dependence of



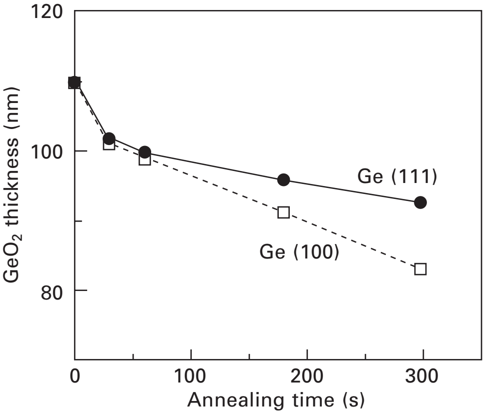
20.12 Substrate orientation effect on electron mobility in Ge(100) and Ge(111) n-MOSFETs at RT. Peak mobility on Ge(111) is about 1100 cm^2/Vs at RT, which is a 1.5-times improvement compared to that of Ge(100). Mobility was estimated by the split-CV method.


Conduction-band energy minima
Equi-energy surface
(100)
(111)
Ge
m_c = 0.153 m_0
m_c = 0.082 m_0 (1 fold)
0.155 m_0 (3 folds)
Si
m_c = 0.19 m_0 (2 folds)
0.32 m_0 (4 folds)
m_c = 0.30 m_0

20.13 Conduction-band energy minima and equi-energy surface in three- and two-dimensional electron systems for Ge and Si, respectively. Electron mobility in Ge MOSFETs on Ge(111) should be higher than on Ge(100) in terms of lower effective mass in the two-dimensional inversion layer.



(a) 1-atm O_2 at 600°C



(b) N_2 at 600°C

20.14 (a) Surface oxide thickness on Ge(100) and (111) grown in 1 atm O_2 at 600°C as a function of oxidation time. This result suggests that Ge(111) should be better in terms of GeO_2 scalability as the interfacial layer of a high-*k* gate stack. (b) Thickness change of GeO_2 films on Ge(100) and (111) wafers by annealing in N_2 at 600°C as a function of annealing time. The initial film thickness was ~110 nm. The GeO volatilization rate from the (111) surface is about half of that from the (100) one. These results strongly support superiority of Ge(111) to Ge(100) in terms of significantly reduced GeO desorption in addition to lower effective mass.

the GeO desorption was investigated. The 110 nm-thick GeO_2 was initially deposited by rf-sputtering. Figure 20.14(b) compares the thickness change of GeO_2 films on Ge(100) and (111) wafers by annealing in N_2 at 600°C as a function of annealing time. The results strongly support the superiority of Ge(111) to Ge (100) in terms of the significantly reduced GeO desorption in addition to the lower effective mass.

To understand the scattering mechanism of μ_{eff}, the temperature dependence of μ_{eff} on Ge(100) and (111) was measured from 150 to 300 K (data not shown). The value of μ_{eff} on both Ge(100) and (111) decreased with measurement temperature, which means that the scattering mechanism at RT was still not governed by phonon scattering. Thus, we conjectured three kinds of possible mechanisms for the mobility degradation.

1. *Ge bulk quality*. Hall effect electron mobility was estimated for n-Ge substrate. Experimentally, the mobility was about 3600 cm^2/Vs for electron, which is comparable to the previously reported electron mobility values of 3900 cm^2/Vs in bulk Ge [16] and sufficiently higher than that of Si. Thus, we concluded that μ_{eff} in Ge n-MOSFET in the present experiment was not dominated by the bulk Ge quality.
2. *Ge/GeO$_2$ interface roughness*. The surface roughness was investigated by AFM. It is generally larger on Ge than on Si, as shown in Table 20.1. Interestingly, the interface roughness of Ge slightly increases after thermal oxidation, while that of Si decreases. This suggests that sacrificial oxidation of Ge might not be appropriate for decreasing roughness, though it is expected that cleaning and etching optimization may hopefully solve this problem. Therefore, in the high N_s region, the electron mobility might be limited by the surface roughness scattering.
3. *Additional Coulomb scattering*. D_{it} near the midgap was low, but the band edge defects as shown in Fig. 20.8 could be a scattering origin for electrons. D_{it} of Ge/GeO$_2$ near the band edge is relatively high compared to that of Si/SiO$_2$, and it might work as an additional Coulomb scattering center. Moreover, most interface states below the Fermi energy are likely to be negatively charged in n-MOSFETs and could work as electron scattering centers in the Coulomb scattering region [17], because the charge neutrality level of Ge is close to the valence band edge, which will be discussed next.

Thus, we expect that further reduction of D_{it} and smoothing of surface roughness on Ge will be the key challenges to achieving much higher electron mobility in Ge n-MOSFETs.

Table 20.1 Comparison of surface roughness between Ge and Si. The present experiment shows a larger surface roughness on Ge than on Si

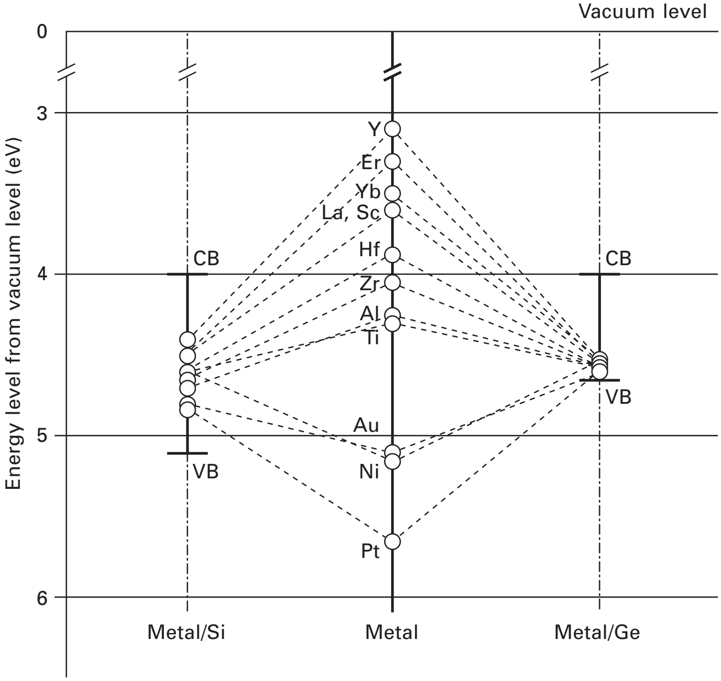
RMS (nm)	Ge(100)	Ge(111)	Si(100)
As-cleaned surface	0.28	0.25	0.15
After oxide removal	0.4	0.35	0.1

20.4 Germanium (Ge)/metal Schottky interface and metal source/drain metal oxide semiconductor field effect transistors (MOSFETs)

20.4.1 Metal/Ge interface

Source/drain technology is another concern of scaled CMOS in terms of enhanced series resistance in MOSFETs. Many efforts have been made for reducing parasitic resistances. Furthermore, heavy impurity doping into Ge for achieving low ohmic resistance at metal/Ge contact is not easy. Therefore, Schottky barrier height (SBH) control is quite important in terms of lowering the contact resistance. On the other hand, the diffused layer resistance also plays a critical role in the total source–drain resistance. If we can use metal in place of a heavily doped semiconductor layer, we will be able to reduce the source–drain resistance significantly. Thus, the metal source–drain junction technology has recently been considered as a solution to the above challenges. To make it feasible, precise SBH control at the metal/Ge interface is indispensable for achieving high-performance CMOS.

The Schottky barrier formation mechanism has been long studied in the history of semiconductor research. The most important and controversial point is how to understand the Fermi-level pinning (FLP) at the metal/semiconductor interface [18]. Ge is a narrow band-gap semiconductor, and strong FLP was reported [19]. To characterize the strong Fermi-level pinning on Ge, we systematically investigated metal/Ge junctions using various metals with a wide range of vacuum work-functions [20]. Figure 20.15 shows the metal work-function of various metals and the energy level of metal Fermi-levels on Si and Ge from the vacuum level, which were estimated from electron affinity and SBH values determined experimentally in *I*–*V* characteristics. Various metals (from Y to Pt) were deposited on Ge by electron-beam evaporation in an ultra-high vacuum chamber. In Fig. 20.15, typical literature values were used for the metal work-function. Note that SBHs for all metals used in the experiments were strongly pinned near the valence band edge of Ge, while moderate FLP was observed in the Si case. Both the pinning level and the pinning parameter were characterized as described in Table 20.2. Those values in Ge are in good agreement with charge neutrality level (branch point in semiconductor) and pinning strength theoretically expected on the basis of metal-induced gap states (MIGS) [21, 22]. This fact does not mean that the strong FLP on Ge is caused by MIGS, but strongly suggests the MIGS mechanism for the Ge case when the narrow band-gap of Ge is considered. Furthermore, it has been demonstrated that the SBH is very little affected by interface modulations, such as forming gas anneal, germanide reaction, or Ge substrate orientation change [20]. This also supports the MIGS mechanism, because the interface defect/disorder formation would be affected by the interface reaction.

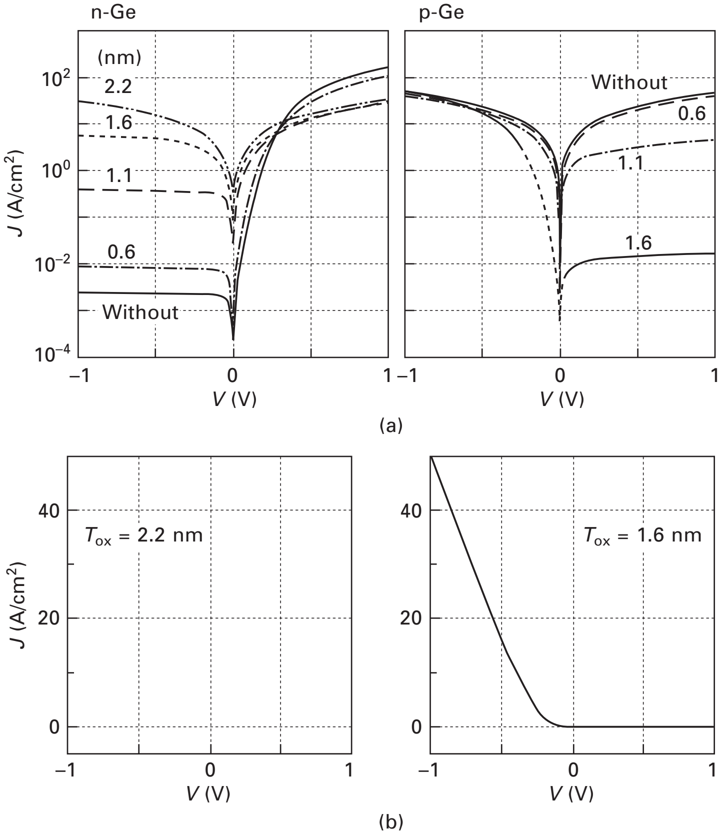


20.15 Fermi-level pinning at both metal/Si and metal/Ge interfaces. In the case of Ge, all metals tested in the experiments are perfectly pinned near the valence band edge of Ge. The effective work-function is defined by the energy difference between the vacuum level and the metal Fermi-level pinned on the semiconductor. It was actually determined by electron affinity of the semiconductor and the Schottky barrier height determined experimentally.

Table 20.2 Comparison of FLP position (charge neutrality level) on Ge from vacuum level and pinning parameter between experimental results and theoretical ones reported in refs [21] and [22]

	Experimental	Calculated [21]	Calculated [22]
FLP position from vacuum level (ev)	4.58	4.48	4.63
Pinning parameter *S*	0.02		0.04

From the viewpoint of metal/semiconductor contact design, it is hard to achieve low ohmic resistance to n-Ge, while it is easy to demonstrate ohmic behavior to p-Ge without any impurity doping. In fact, ohmic characteristics for metal/p-Ge and Schottky characteristics for metal/n-Ge were experimentally observed irrespective of the metal work-function in a vacuum. However, we paid attention to the fact that the flat-band voltage (V_{FB}) and surface potential in Ge MOS capacitors were substantially modulated without showing strong FLP at the metal/insulator and/or the insulator/Ge interface. This implies that the strong FLP observed at metal/Ge junctions seems to be unpinned by inserting a thick oxide between the metal and Ge. However, the thick oxide insertion cannot be used for source/drain junctions. Therefore, we investigated the effect of ultra-thin oxide (UTO) insertion into a metal/Ge interface. Figure 20.16(a) shows the *I–V* characteristics of metal/Ge contacts in semi-log scale with and without an ultra-thin GeO_2 layer between them [23]. The linear scale *I–V* characteristics with UTO are also shown in Fig. 20.16(b). Clear transitions from ohmic to Schottky in p-Ge and from Schottky to ohmic in n-Ge are observed. Although a modulation of SBH even in Si has been observed [24], drastic transitions between ohmic and Schottky for



(a)

(b)

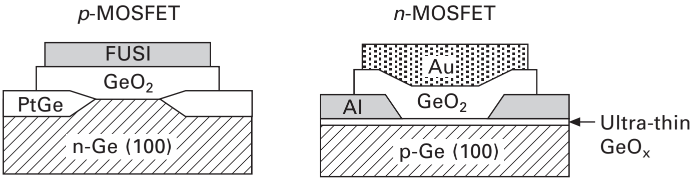
20.16 Ultra-thin oxide thickness dependence of *J–V* characteristics at the metal/UTO/Ge junction, where ultra-thin oxide (UTO) was GeO_x. SBH gradually decreased on n-Ge and increased on p-Ge. *J–V* characteristics are an a semi-log scale in (a), while those in the case of T_{ox} = 1.6 nm are plotted on a linear scale in (b).

both n- and p-Si have never been reported. This method enables us to achieve metal source/drain n-channel FETs without any impurity doping into Ge. It is not inconsistent qualitatively with the MIGS model for strong FLP at the metal/Ge interface, because FLP due to the MIGS should be affected by suppressing the metal wave function tailing into Ge. However, we should notice that the pinning parameter is not recovered to the Schottky limit ($S = 1$) by inserting UTO, which means other factors should be involved in the strong FLP at metal/Ge interface as well [25]. Recently other groups reported the UTO insertion effect by changing the insulator on Ge [26]. Further investigation is still needed to unquestionably understand the FLP mechanism at metal/Ge interfaces.

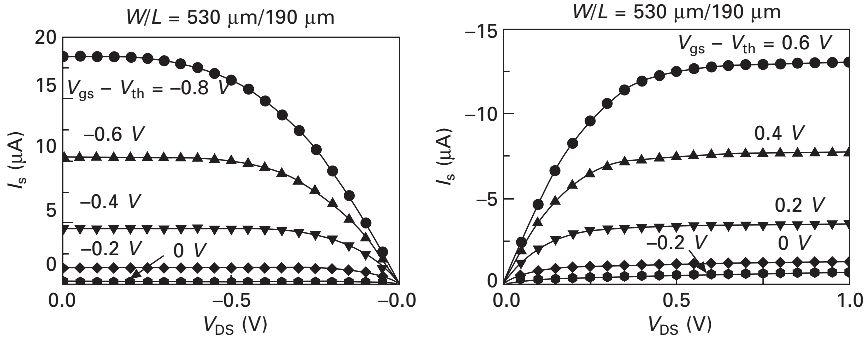
20.4.2 Metal source/drain MOSFETs

We applied this technique to the fabrication of metal source/drain Ge n-MOSFETs without any impurity doping. Figure 20.17 shows schematics of the fabricated Al source/drain Ge p- and n-MOSFETs. PtGe was formed on n-Ge for p-MOSFET, while about 2 nm-thick GeO_x was formed on p-Ge substrates, followed by the thermal evaporation of Al for the source/drain electrodes. Then GeO_2 was deposited as the gate insulator, and Au was deposited for the gate electrode for n-MOSFET, while NiSi FUSI was employed for p-MOSFET [27].

Figure 20.18 shows output *I–V* characteristics of n- and p-MOSFETs with GeO_2 gate dielectric films that were made without using ion implantation. The peak mobility values of electrons and holes without any correction were 270 and 370 cm^2/Vs, respectively, though no process optimization was employed for either gate stack or junctions. The results in n-MOSFET clearly show that electrons are successfully injected from the Al source to the inverted electron channel thanks to the ultra-thin GeO_x film insertion, and that the channel conduction is well controlled by the gate bias. The output characteristics in the low V_{DS} region indicate a very low parasitic resistance in spite of no impurity doping in the source/drain regions. Considering the low solubility limit and large diffusion coefficient of n-type impurities in Ge,



20.17 Schematic views of metal source/drain Ge p- and n-MOSFETs. In the case of n-MOSFET, ultra-thin GeO_x is inserted between source/drain metal and p-Ge.



20.18 I_S–V_{DS} characteristics of fabricated metal source/drain Ge (100) p- and n-MOSFETs. Peak mobilities of 370 cm^2/Vs for (a) p-MOSFET and 270 cm^2/Vs for (b)n-MOSFET were obtained with no series resistance correction. In this case, GeO$_2$ was deposited for gate dielectric film by rf-sputtering. Thus, mobility values were lower than the results shown in Fig. 20.11, because both the interface and bulk properties of the Ge/GeO$_2$ system were worse than those grown by HPO + LOA.

it is quite likely that ultra-short channel Ge n-MOSFETs might be fabricated with the metal source/drain structure.

This experimental proof is quite convincing in considering the metal source/drain Ge n-MOSFETs, and will open a promising way for the metal/source drain Ge CMOS.

20.5 Conclusion and future trends

In this chapter we have focused on both Ge/GeO$_2$ and Ge/metal interfaces from the viewpoint of electronic states formed at those interfaces. Those two interfaces are essentially important for building up MOSFET structure.

The value of D_{it} at the Ge/GeO$_2$ interface was demonstrated to be of the order of 10^{10} cm^{-2} eV^{-1} at the mid-gap. The very low value of D_{it} is a little surprising, considering the fact that the low D_{it} has been achieved without using any intentional technique for passivating dangling bonds possibly located at the Ge/GeO$_2$ interface. The passivation process is quite different than in the Si/SiO$_2$ case and the low D_{it} might be attributable to more flexible bonding in Ge–O than in Si–O. Thus, we have regarded it as the self-passivation process in the Ge/GeO$_2$ system.

The results were applied to n-MOSFET. We demonstrated the highest electron mobility reported so far in Ge FETs. We think this result is quite important from the viewpoint of the controllability of the interface. This was achieved by thermodynamic and kinetic control of the Ge/GeO$_2$ stack. In addition, Ge(111) was revealed to be much better than Ge(100) from the viewpoint of thermal process stability as well as lower effective mass. As for the electron mobility degradation mechanism, significantly large scattering by charged interface states only for electrons was discussed from the viewpoint of the charged neutrality level position of Ge, in addition to the surface roughness scattering. Those challenges must be overcome from materials science viewpoints for further high-performance Ge MOSFETs. Ge n-MOSFETs are not intrinsically degraded but are quite promising in terms of their high mobility, which is much higher than the Si universal mobility.

Ge/metal interfaces showed strong Fermi-level pinning, and we found that insertion of an ultra-thin insulator between the metal and Ge could modulate the Schottky barrier height. By using this fact, we demonstrated quite a good operation of the metal source/drain n-MOSFET as well as p-MOSFET experimentally. Although further optimization is definitely needed for achieving high performance Ge CMOS, metal source/drain technology should play a more significant role in reducing the series resistance effect relatively in smaller MOSFETs. It might also provide us a key suggestion for elucidating the Schottky barrier formation mechanism that has been discussed for many years in the semiconductor interface community.

In fact, there remain many challenges such as scalability issues, including high-*k* gate stack technology and short channel effect immunity, in Ge CMOS, but we believe that Ge CMOS is quite promising for putting higher performance semiconductor devices beyond the current Si-CMOS into practice.

20.6 Acknowledgments

The author would like to thank his colleagues, K. Kita, T. Nishimura and K. Nagashio, and graduate students, S. K. Wang, C. H. Lee and M. Yoshida. This work was partly supported by a Grant-in-Aid for Scientific Research (S) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and was partly performed in collaboration with STARC.

20.7 References

1. D. A. Antoniadis and A. Khakifirooz, *Tech. Dig. IEDM* (2007) 257.
2. H. Shang, M. M. Frank, E. P. Gusev. J. O. Chu, S. W. Bedell, K. W. Guarini, and M. Ieong, *IBM J. Res. Dev.* **50** (2006) 377.
3. A. Toriumi, T. Tabata, C. H. Lee, T. Nishimura, K. Kita, and K. Nagashio, *Microelectronic Engineering* **86** (2009) 1571.
4. A. Toriumi, S. K. Wang, C. H. Lee, M. Yoshida, K. Kita, T. Nishimura, and K. Nagashio, *ECS Trans.* **28**(2) (2010) 171.
5. E. P. Gusev, H. C. Lu, T. Gustafsson, and E. Garfunkel, *Phys. Rev.* **B 52** (1995) 1759.

6. K. Kita, S. Suzuki, H. Nomura, T. Takahashi, T. Nishimura, and A. Toriumi, *Jpn. J. Appl. Phys.* **47** (2008) 2349.
7. S. K. Wang, K. Kita, C. H. Lee, T. Tabata, K. Nagashio, T. Nishimura, and A. Toriumi, *Ext. Abst. SSDM* (2009) 1002; *J. Appl. Phys.* **108** (2010) 054104.
8. K. Kita, S. K. Wang, M. Yoshida, C. H. Lee, K. Nagashio, T. Nishimura, and A. Toriumi, *Tech. Dig. IEDM* (2009) 693.
9. K. Nagashio, C. H. Lee, T. Nishimura, K. Kita, and A. Toriumi, *MRS Symp. Proc.* **1155** (2009) C06-02.
10. C. H. Lee, T. Tabata, T. Nishimura, K. Nagashio, K. Kita, and A. Toriumi, *Appl. Phys. Express* **2** (2009) 071404.
11. M. Yoshida, K. Kita, K. Nagashio, T. Nishimura, and A. Toriumi, Abstract, *IEEE SISC* (2009) 28.
12. H. Hosono, Y. Abe, D. L. Kinser, R. A. Weeks, K. Muta, and H. Kawazoe, *Phys. Rev.* **B 46** (1992) 11445.
13. C. H. Lee, T. Nishimura, N. Saido, K. Nagashio, K. Kita, and A. Toriumi, *Tech. Dig. IEDM* (2009) 457.
14. K. Kita, T. Takahashi, H. Nomura, S. Suzuki, T. Nishimura, and A. Toriumi, *Appl. Surf. Sci.* **254** (2008) 6100.
15. S. Takagi, *VLSI Symp.* (2003) 115.
16. W. C. Dunlap Jr, *Phys. Rev.* **77** (1950) A759.
17. P. Tsipas and A. Dimoulas, *Appl. Phys. Lett.* **94** (2009) 012114.
18. W. Mönch, *Semiconductor Surfaces and Interfaces*, 3rd edn., Springer-Verlag (2001).
19. A. Dimoulas, P. Tsipas, A. Sotiropoulos, and E. K. Evangelou, *Appl. Phys. Lett.* **89** (2006) 252110.
20. T. Nishimura, K. Kita, and A. Toriumi, *Appl. Phys. Lett.* **91** (2007) 123123.
21. J. Tersoff, *J. Vac. Sci. Technol.* **B4** (1986) 1066.
22. M. Cardona and N. Christensen, *Phys. Rev.* **B35** (1987) 6182.
23. T. Nishimura, K. Kita, and A. Toriumi, *Appl. Phys. Express* **1** (2008) 051406.
24. D. Connelly, C. Faulkner, P. A. Clifton, and D. E. Grupp, *Appl. Phys. Lett.* **88** (2006) 012105.
25. T. Nishimura and A. Toriumi, *unpublished*.
26. R. R. Lieten, S. Degroote, M. Kujik, and G. Borghs, *Appl. Phys. Lett.* **92** (2008) 022106; Y. Zhou, M. Ogawa, X. Han, and K. L. Wang, *Appl. Phys. Lett.* **93** (2008) 202105; M. Kobayashi, A. Kinoshita, K. Saraswat, H.-S. P. Wong, and Y. Nishi, *J. Appl. Phys.* **105** (2009) 023702.
27. T. Takahashi, T. Nishimura, L. Chen, S. Sakata, K. Kita, and A. Toriumi, *Tech. Dig. IEDM* (2007) 697.

21

Silicon (Si) and germanium (Ge) in optical devices

K. OHASHI, NEC Corporation, Japan

Abstract: This chapter describes optical devices which use Si, SiGe, and Ge technologies for electronics. It starts with the optical waveguides and optical resonators which create slow light for microphotonic devices. Then, electro-optical modulation, photo-carrier generation, and photo emission will be discussed on the basis of nano- and micro-structures. Both modulators and detectors are critical components of the on-chip optical interconnections that are expected to enable low-power-dissipation electronics systems. Ultra-small electric capacitance accomplished through silicon photonics and nano-photonics would provide a strong tool for that purpose. Multi-modal sensing is also a promising area for these new optical devices.

Key words: photodiode, electro-optical modulator, silicon photonics, surface-plasmons, optical waveguide.

21.1 Background

High-speed optical communication may go back to the use of smoke signals along the Great Wall to alert soldiers of enemy attacks in ancient China. Optical device technology has been supplying a fundamental infrastructure for telecommunication. Optical devices are also key components for imaging and sensing in the digital world, and they will therefore be indispensable in the ubiquitous network society.

Nevertheless, they have a drawback of being fairly large compared with electronic ones. Integrated photonic circuits have been investigated almost since semiconductor integrated circuits were announced (Anderson 1965); however, their level of integration is still low. The fundamental difference in size between optical and electronic components results from the fact that the wavelength of light is longer than the de Broglie wavelength of electrons. There are two approaches to closing this gap in the difference in device size. The first is to slow down the group velocity of light in photonic components. The second is to use near-field optics to reduce the size of photonic devices beyond the diffraction limit.

While the performance of electronics systems has been improved mainly due to the miniaturization of transistor gate switches, reducing the width of metal interconnects increases their resistance and capacitance, thereby increasing latency and power dissipation (Meindl *et al.* 2002). More than 50% of the dynamic power in recent large-scale integrated circuits (LSIs) has been consumed by interconnections (Magen *et al.* 2004) and this percentage is still increasing. This brought about the idea of introducing the revolution to interconnections, shifting optical communication from telecom to on-board and on-chip interconnection. The fundamental drawback with electrical interconnects is that they require the electrical capacitance of electrical lines to be charged and discharged quickly when data are transported on the lines. The delay in data transmission by interconnections is usually expressed by RC where R is the resistance and C is the capacitance of the wires. Since both R and C are proportional to the length of the wire, the delay increases as the square of the wire's length. To avoid this rapid increase, global-level interconnections used in LSIs, where long copper wires (some of which are longer than several millimeters) connect functional blocks, employ repeater circuits to divide the wire into shorter sections, thereby reducing delay at the cost of the complexity of the system on a chip (Bakoglu 1990).

Optical interconnections (Miller 2000, 2009), on the other hand, require an optically transparent medium. Power dissipation in optical interconnects mainly results from optical loss due to light-scattering, which can be considerably lower than the loss from the resistance of electric wires. The intrinsic issue with power dissipation results from the conversion of electrical signals to optical ones and vice versa (Ohashi *et al.* 2009). Imaging and sensing have fundamentally been achieved with analog technology. These have recently been gradually fitted into digital electronics. Silicon photonics in the broad sense of the term, including imaging, sensing and solar cell applications, will also be reviewed.

21.2 Optical waveguides

Single-mode optical fibers in long-haul communication systems have become established as an extremely low-loss medium technology for optical wavelength-division-multiplexing (WDM) data transmission (Yariv and Yeh 2007). They usually use silica (SiO_2) fibers that consist of a core and surrounding cladding, where they have been lightly doped with germania (GeO_2) (Friebele *et al.* 1984), for example, to raise the refractive index of the core to slightly more than that of the cladding (Kao and Davies 1968). This structure confines the signal light mainly in the core by reflecting the light at the interface between the core and the cladding. The main role played by the cladding is to prevent contaminated particles from adhering to the surface of the core. Typical values of the refractive index are 1.455 for the cladding and 1.47 for the core. This small difference is important for long-haul communication because the optical loss at the interface is small.

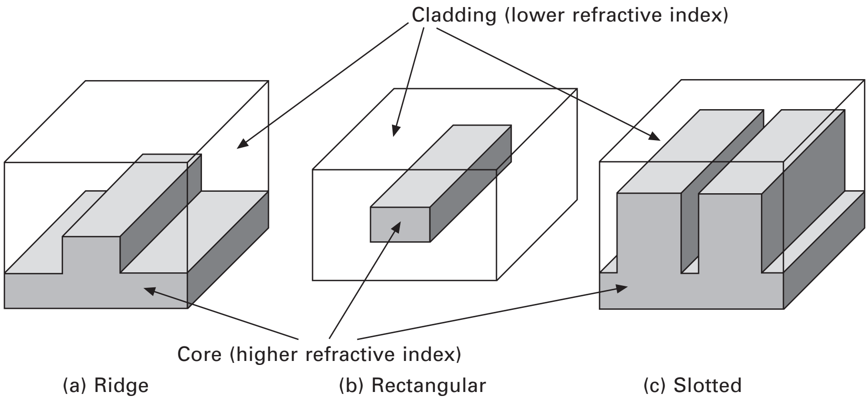
However, it prevents optics from being introduced to small on-chip scale interconnections.

21.2.1 Waveguide structures

The structures of optical waveguides are fundamentally similar to those of optical fibers except that the waveguides do not have circular symmetry. They can be classified into four types:

1. *Slab (or planar) waveguides*. These are two-dimensional waveguides. A core layer is sandwiched by a substrate and cladding layers with a slightly lower refractive index.
2. *Ridge (or rib) waveguides*. These have a thicker portion along the direction of light propagation in the core layer of slab waveguides (Fig. 21.1(a)).
3. *Rectangular waveguides*. These have a core with a square or rectangular cross-section (Fig. 21.1(b)).
4. *Wavelength and sub-wavelength-scale structures*. These structures in waveguides provide several useful optical functions to optical devices and circuits.

Slab waveguides do not change the direction of light propagation. Rectangular waveguides confine the light more strongly than ridge (or rib) waveguides do, so that a smaller bending radius is obtained with rectangular waveguides. Nano-structures can be combined with the others of types (1)–(3) to provide slow light propagation. We can derive both transverse electric (TE) and transverse magnetic (TM) modes of light propagation in planar waveguides, where TE-modes have their electric field lies in the plane and perpendicular to the propagation direction, and TM-modes have their


Cladding (lower refractive index)
Core (higher refractive index)

(a) Ridge (b) Rectangular (c) Slotted

21.1 Optical waveguide structures.

magnetic field in the same direction (Reed and Knights 2004). Similarly, we can calculate orthogonal TE-like and TM-like modes for more complicated waveguide structures.

21.2.2 Si and SiON for waveguides

There are two kinds of materials employed in Si-based optical waveguides. The first is semiconductor Si for wavelengths longer than 1200 nm, which corresponds to a band gap of 1.12 eV (1.14 eV at room temperature (RT)), and it is therefore transparent with near-infrared light for telecommunication (wavelength 1300–1600 nm) (Soref and Lorenzo 1985). The second kind is dielectric materials such as silicon oxides and silicon nitrides. The band gap of these materials is large enough (e.g. 9.0 with SiO_2) to cover visible light (wavelength 390–700 nm).

The absorption mechanism for photons at wavelengths longer than 1200 nm in Si is free-carrier absorption, which can be used for the electro-optic modulators discussed in the next section (21.3), although absorption in Si is usually negligible unless the concentration of carriers is increased by doping.

Higher-index-contrast waveguides have been developed for short-distance optical interconnections to achieve sharper bending of optical waveguides than silica fiber. Si waveguides, where the core material is Si and the cladding material is silica, have been developed for such applications. Since the refractive index, n, of Si (for infrared light with a wavelength of ~1550 nm) is as large as ~3.5, once light enters the Si core, it is strongly confined in the core and a minimum bending radius for the waveguide of less than 2 μm is achieved. The propagation loss of Si waveguides is ~1 dB/cm (Tsuchizawa *et al.* 2006, Gnan *et al.* 2008), which is acceptable in on-chip optical communications. Eliminating the etching process to reduce side-wall roughness decreased the propagation loss down to 0.3 dB/cm at a cost of slightly large bending radius of 50 μm (Cardenas *et al.* 2009), which might be overcome by introducing a mirror for bending (Qian *et al.* 2006).

Silicon oxynitride (SiON) and silicon nitride (Si_3N_4) are dielectric materials with high transparency that can be used for the core, while SiO_2 can be used for the cladding to construct relatively high-contrast-index waveguides. The refractive index of SiON systems can be adjusted and varies from 1.45 (SiO_2) to 2.0 (Si_3N_4) according to the content of O and N (Valetta *et al.* 1989, de Ridder *et al.* 1998). This feature provides large flexibility in optical waveguide design. A SiON waveguide with a core width of 2 μm indicated the insertion loss of 0.2–0.3 dB/cm at a wavelength of 850 nm in the TM mode (Tsuchizawa *et al.* 2008).

21.2.3 Wavelength and sub-wavelength structures

A grating causes periodic modulation of the effective refractive index. The period corresponds to the effective wavelength of light at the grating surface. Gratings are sometimes used to couple external light to an optical circuit by controlling the diffraction of the light (Dakss *et al.* 1970, Tamir 1975). They are also used as a Bragg reflector or Fabry-Pérot filter to reflect or transmit the specific wavelength of light one-dimensionally (Yariv and Yeh 2007). Photonic crystals are a kind of metamaterial with spatial periodicity that create a frequency window or photonic band gap in which the propagation of light is inhibited (Ohtaka 1979, Yablonovitch 1987). Defects in periodicity can locally destroy the band gap and confine light at the defects (John 1987, Sigalas *et al.* 2001, Novotony and Hecht, 2006). This technology is employed for slow-light technology, which is useful for obtaining optical devices, especially those for electro-optical and opto-electric conversion to reduce electric capacitance (Baba 2008). Three-dimensional (3-D) photonic crystals have a photonic band gap along all directions, and those with a 2-D periodic structure may have a photonic band gap along all directions in a plane (Yablonovitch 1987).

Bragg reflectors and Fabry-Pérot interferometers are classified as 1-D photonic crystals. Fabry-Pérot interferometers, which consist of plane-parallel dielectric films or metal films, are cavities that operate as optical resonators. Confined light creates strong optical intensity in the cavities. An optical resonator with curved mirrors confines light two-dimensionally as well as one-dimensionally along the direction that light is propagating. The extent of confinement is of the order of the wavelength. Photonic crystal technology may affect the development of a small photodiode and a small laser light source, as will be discussed later.

The photonic crystal sharply bends the propagation direction of light at a specific angle. The layout of the periodic structure, square or triangular, greatly changes the band structure. It also produces an optical cavity with an extremely high Q-value and slow-light mechanism with a high effective refractive index. Two-dimensional Si photonic crystal using a silicon-on-insulator (SOI) wafer has at least two different structures. The first is an Si slab with periodically aligned holes (Lončar *et al.* 2000). The second is periodically aligned Si pillars (Tokushima *et al.* 2004). A square-lattice-of-rods photonic crystal has a line defect, which acts as a waveguide and has an advantage in that it couples with an Si waveguide.

Nano-slotted waveguides (Xu *et al.* 2004), where strongly confined light is guided in a subwavelength-scale slot with a lower refractive index, which is sandwiched between silicon strips or slabs, are also useful for such applications with further degrees of freedom to bend the angle of the waveguide (Fig. 21.1(c)). Slotted Si waveguides and pillar-type Si photonic crystals are suitable for obtaining active optical devices by filling functional materials such as electro-optical polymers in the space between the silicon. A Si-based hollow-core surrounded by SiGe/Si multi-layer claddings is proposed for the long-wave infrared region from $\lambda = 1.2–100\ \mu m$ (Soref *et al*. 2006).

Optical resonators can be constructed using optical waveguides. Ring resonators, which include modified versions of racetrack resonators and disk resonators, are typical optical resonators usually coupled with an external waveguide. The optical filters for WDM technology can be fabricated fairly compactly (less than several tens of micrometers) using multiple rings with different resonant wavelengths. In some cases, the quality factor (Q-factor) of such optical resonators is too high for them to operate under harsh environments such as those involving rapid temperature changes on an LSI chip.

21.3 Modulators

Electro-optic (EO) modulators convert electrical signals to optical signals. Optical switches can be constructed using the same scheme when the on–off ratio is high enough, whereas modulators usually require low-voltage operation. Examples of high-speed (>1 GHz) EO materials are ceramics such as lithium niobate ($LiNbO_3$) and polymers with a chromophore dispersed in a polycarbonate. Liquid crystals can be applied with silicon LSI for low-speed (~100 Hz) modulation (Cotter *et al.* 1990). The refractive index of electro-optic materials changes by applying an electric field. The linear electro-optic effect (Pockels effect) is useful for EO modulators that are driven by relatively low-voltage signals. The second electro-optic effect (Kerr effect), on the other hand, is useful for optical switches that require large contrast in the signal intensity of the on- and off-states. Unfortunately, the linear electro-optic effect of Si is fundamentally zero due to its crystallographic centro-symmetry, while its second-order electric field effect changes the refractive index so that it is as small as 10^{-4} at more than the breakdown electric field intensity (10^6 V/cm) (Soref and Bennet 1987). Therefore, we have three ways of developing silicon-based photonic devices. The first is to break the crystallographic symmetry. Jacobsen *et al.* (2006) reported that a significant linear EO effect is induced by depositing a straining layer on an Si waveguide. The second is to use physical mechanisms other than the EO effect such as free-carrier absorption and the thermo-optical (TO) effect. The third is to use non-Si materials, which can be fabricated onto an Si wafer.

21.3.1 Free carrier absorption

The use of electro-optic modulators using free-carrier absorption is one of the most active areas in Si photonics (Soref and Bennet 1987). The refractive index of semiconductors changes with their concentration of free carriers. The fast response of free carrier density in silicon waveguides can be used to modulate their refractive index. Optical modulators employing this mechanism are called Si-waveguide-type modulators. These modulators can be categorized into two types: (1) carrier injection, and (2) carrier depletion. The former injects free carriers to cause plasma effects, the latter changes the carrier density near the interface of MOS capacitors.

Carrier-injection modulators have limits in high-speed operation due to their relatively slow response originating from the recombination of carriers. Nevertheless, gigahertz operation is achieved by employing a pre-emphasis technique to input electric signals to inject and remove carriers within a short time (Jongthanmmanurak *et al.* 2006). Green *et al.* (2007) reported 10 Gb/s, 5 pJ/bit data transmission of a carrier-injection-type optical modulator 100–200 μm-long using a Mach–Zehnder interferometer (MZI) which converts the relative phase shift between two collimated light beams from a coherent light source to the amplitude of optical output.

Carrier-depletion modulators change the thickness of the depletion layer near the electrodes by changing the voltage between them. This electro-optic modulation usually operates faster than that of carrier-injection modulators and does not need a pre-emphasis technique (Liu *et al.* 2008).

21.3.2 Thermo-optical effect

The refractive index of waveguide material can be modulated by changing the temperature as well as applying an electric field. The thermo-optic coefficient, $\mathrm{d}n/\mathrm{d}T$, of Si is $\sim 1.86 \times 10^{-4}$, which is higher than that of SiO_2, $\sim 1 \times 10^{-4}$. Combining this effect with a compact Si optical circuit makes it possible to attain a thermo-optic modulator and a thermo-optic switch for relatively low-speed (up to 1 MHz) applications (Cocorullo *et al.* 1995, Espinola *et al.* 2003, Chu *et al.* 2005, Yamada *et al.* 2006).

21.3.3 Resonant structure for modulators

Resonance structures such as ring-type waveguides could be effectively introduced. The size of the resonators could be less than 10 μm (Lipson 2006). A small carrier-injection-type modulator with a ring-type waveguide has been reported to operate at 18 Gb/s (Manipatruni *et al.* 2007). Another interesting resonator for modulators is photonic crystal. Wülbern *et al.* (2009) demonstrated an electro-optic modulator with a two-dimensional photonic-crystal slab of Si for the core material and an electro-optic polymer for that of the cladding. The expected switching voltage was sufficiently low (~1 V) for high-speed modulation (100 GHz).

Due to the relatively large TO effect, optical circuits using an Si waveguide with a resonant structure are sensitive to temperature fluctuations, which are inevitable in on-chip applications. To overcome this problem, a composite waveguide with a polymer, which has a negative thermo-optic coefficient (Han and Wang 2007), and local heating of Si waveguide have been investigated (Manipatruni *et al.* 2008). Coupled resonator optical waveguides (CROWs) provide flat-top band-pass filters that are robust against temperature changes obtained by serially coupling several resonators with slightly different resonant frequencies (Vlasov *et al.* 2008). Guha *et al.* (2010) proposed and demonstrated a ring resonator combined with an MZI, which was passively temperature compensated.

21.3.4 Electro-absorption

Unlike the EO effect, the electro-absorption (EA) effect results from the distortion of energy bands caused by the applied electric field. There are two types of EA effects in semiconductors: the Franz–Keldysh effect for bulk semiconductors and the quantum-confined Stark effect (QCSE) for quantum wells (Miller 2009). The latter involves steeper changes in the spectrum due to discretization of the density of states and the stronger excitonic effects in quantum-confined structures, and therefore the absorption coefficient changes more abruptly than in the former.

EA optical modulators using Ge quantum wells can be made as small as 10 μm and these are expected to operate at up to 500 Gb/s without resonators (Miller 2009). A modulator using the QCSE of a Ge quantum well at $\lambda = 1440–1460$ nm has been reported (Kuo *et al.* 2005, Roth *et al.* 2007, Schaevitz *et al.* 2008). Another group reported a modulator using the Franz–Keldysh effect of strained Si on Ge to operate at $\lambda = 1610–1640$ nm (Jongthanmmanurak *et al.* 2006), and strained GeSi at $\lambda = 1539–1560$ nm with an estimated extremely low energy consumption of 50 fJ/bit (Liu *et al.* 2008).

21.3.5 Other high-speed electro-optic materials

Hetero-devices using *c*-axis oriented ferroelectric lithium niobate ($LiNbO_3$) films grown on SiO_2 buffer layers on Si substrates have been developed for compact blue-green lasers (Lee and Feigelson 1999). EO polymers have also been studied for EO modulation (Tomme *et al.* 1991). As the electro-optic effect in polymers is caused by the displacement of electrons, not ions, a fast response at more than 100 GHz can be obtained (Lee *et al.* 2000). Brosi *et al.* (2008) infiltrated an EO polymer. The non-epitaxial growth of nano-crystalline EO ceramic has been studied for on-chip modulators and optical switches. A lanthanum-modified lead-zirconate-titanate (PLZT) film was non-epitaxially deposited on an Si wafer by using aerosol-deposition.

The average grain size of the film was about 20 nm, which is much smaller than the wavelength, and it was, therefore, highly transparent (Nakada *et al*. 2004). The EO coefficient of the film became ~100 pm/V after the films were annealed at 600°C. An EO modulator fabricated on an Si wafer using this method operated in the GHz range (Shimizu *et al*. 2009).

21.4 Photodetectors and photovoltaics

The wavelength band available for Si photodiodes is basically less than 1.1 μm because the band gap of Si is 1.14 eV at RT. Since Si is an indirect band-gap semiconductor, the absorption length of photons is fairly long (Donati 1999, Rieke 2003). The absorption length of Si is ~2 μm at $\lambda = 0.6$ μm, ~10 μm at $\lambda = 0.8$ μm, and ~40 μm at $\lambda = 1.0$ μm. The Si photodiodes for red or infrared light are therefore relatively large and their response is usually not fast (<100 MHz). Germanium has instead been extensively investigated as a high-speed detector for optical interconnections since its absorption length is ~1 μm even at $\lambda = 1.3$ μm. When the light intensity is strong enough, however, two-photon absorption in Si can be observed at the 1.5 μm wavelength (Tsang *et al*. 2002, Bristow *et al*. 2007). For longer wavelength applications, silicon bolometers are used instead of photodiodes.

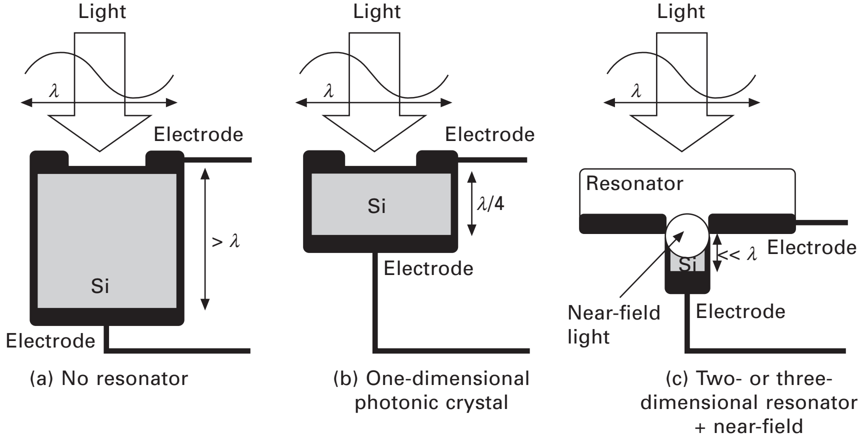
21.4.1 Si photodiodes

The two main factors limiting the response time of photodiodes are the transit time of photogenerated carriers and the electrical capacitance of photodiodes. To reduce both factors, incident light should be efficiently confined within a small active region.

The optical power of the light source in the photodiode circuitry that is required to obtain sufficient signal voltage for an electronic circuit in the subsequent stage is proportional to the sum of the electric capacitance of the photodiode C_{PD} and the input CMOS capacitance C_{Load} (Ohashi and Fujikata 2009) at high frequency. In the latest CMOS technology, the C_{Load} of a transistor with which a detector is integrated is ~2 fF for the 90-nm CMOS-technology node (Ohashi *et al*. 2006), and it will be 0.38 fF for the 32 nm node (Miller 2009). In contrast to this, as C_{PD} is 100 fF or more, it is essential to reduce C_{PD} to enable the total system including the light source to operate at low power.

Si is an indirect band gap material and it has not been used in high-speed photodiodes. Its optical absorption length is about 10 μm when wavelength λ is ~800 nm, which is more than ten times longer than that of Ge. Conventional Si photodiodes, therefore, have a large volume to gain enough sensitivity as shown in Fig. 21.2 (a).

To reduce the transit time of carriers by decreasing the volume of Si,


Light
λ
Electrode
> λ
Si
Electrode
(a) No resonator
Light
λ
Electrode
Si
λ/4
Electrode
(b) One-dimensional photonic crystal
Light
λ
Resonator
<< λ
Electrode
Si
Near-field light
Electrode
(c) Two- or three-dimensional resonator + near-field

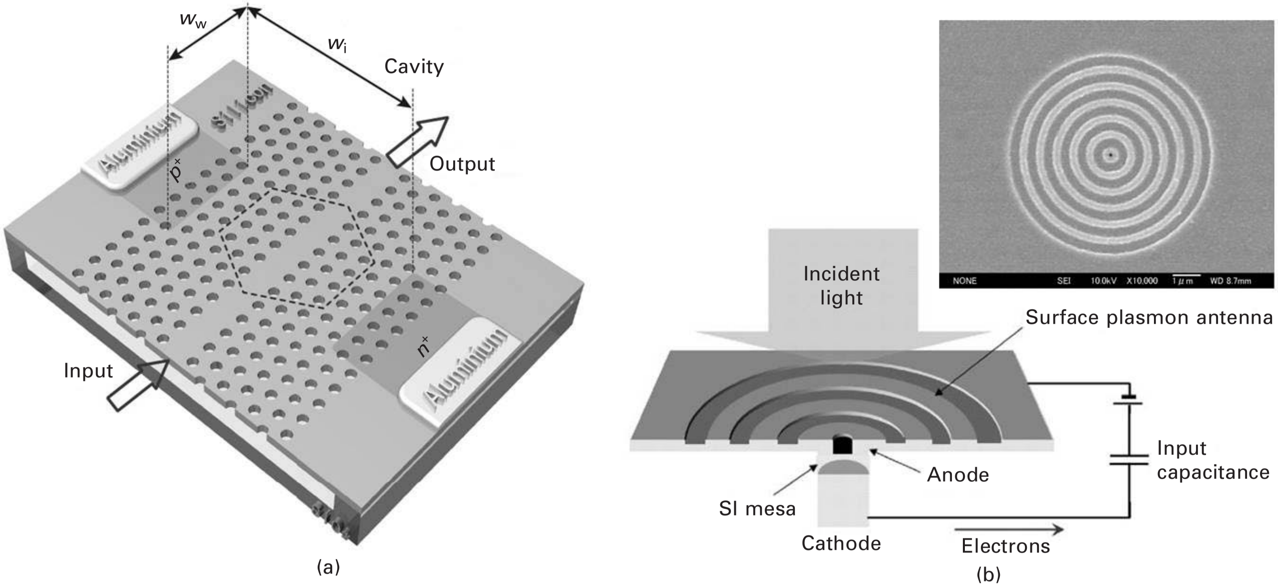
21.2 Confinement of light by resonators in small active area.

distributed Bragg reflectors (DBRs) or 1-D photonic crystals have been applied to Si photodiodes (Fig. 21.2(b)) (Schaub *et al.* 1999, Emsley *et al.* 2002), although the DBR is high-Q and the resonant wavelength is sensitive to temperature. Confinement of light using 2-D or 3-D resonators further reduces the size of Si photodiodes (Fig. 21.2(c)). Reduction of electrode area gives small photodiode capacitance C_{PD} as well as short transit time.

Stronger confinement is achieved by introducing 2-D photonic crystal structures (Fig. 21.3(a)). A p-i-n structure was integrated into a photonic crystal nano-cavity to confine light in an Si photodiode (Fig. 21.3(a)). This photodiode operated in the 1550-nm band, which can be explained by two-photon absorption (Tanabe *et al.* 2010). Chen *et al.* (2009) also reported a cavity-enhanced photocurrent generation in the 1550-nm band.

Ishi *et al.* (2005) proposed a silicon nano-photodiode, in which a semiconductor mesa structure was located just under a sub-wavelength aperture of a surface-plasmon antenna shown in Fig. 21.3(b). The mesa could be as small as the thickness of the Schottky depletion layer. Incident light in the nano-photodiode excites surface plasmons on the antenna, and they enhance the near-field on the exit side of the aperture. Then, the antenna generates photocarriers in the depletion layer. As the separation between electrodes can be reduced to as small as the size of the near-field, the transit time of the carriers can be reduced to several picoseconds even when the semiconductor material is an indirect semiconductor such as Si. The built-in potential of the Schottky barrier permitted low-bias operation of the nano-photodiode at high frequency (Fujikata *et al.* 2005).

A silicon nano-photodiode coupled with an SiON waveguide has been developed for an on-chip optical interconnect (Fujikata *et al.* 2008). Periodic Ag metal–semiconductor–metal (MSM) electrodes at the interface of the


w_w
w_i
Cavity
Output
Aluminium
Silicon
p^+
n^+
Aluminium
Input
(a)
NONE SEI 10.0kV X10,000 1 μm WD 8.7mm
Incident light
Surface plasmon antenna
Input capacitance
Anode
SI mesa
Cathode
Electrons
(b)

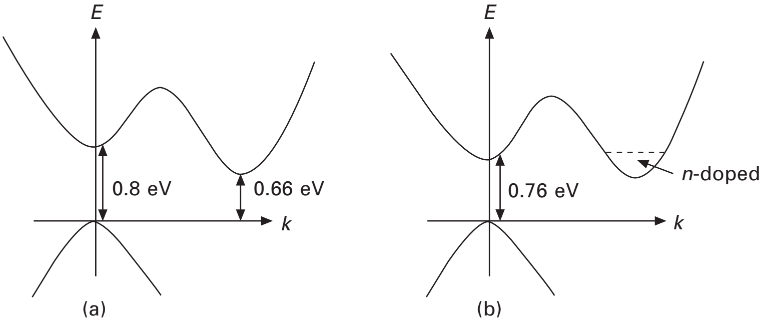
21.3 Light confinement structures for silicon photodiode: (a) photonic crystal waveguide (reproduced from Tanabe *et al*. (2010) with permission), and (b) surface-plasmon antenna.

waveguide and Si layer diffracted the propagating light to obtain a short coupling length of ~10 μm. A waveguide-integrated silicon nano-photodiode to demonstrate on-chip optical clock distribution was fabricated on an n-type silicon-on-insulator (SOI) wafer.

A surface-plasmon antenna in a nano-photodiode acts as a resonant cavity that temporarily reserves light energy and generates photocarriers. Its function as a cavity is similar to that of Bragg reflectors and photonic crystals. The cavity function makes sense when resonance is achieved within a shorter time than the response time of the photodiode. Figure 21.2(c) shows a basic concept of the light confinement in photodiodes. While 1-D photonic (DBRs) crystals reduce the transit time of photocarriers, small near-field light generated from 2-D or 3-D resonators (photonic crystals and surface-plasmon antennas) further reduces the electrode area to give small electric capacitance which is essential to integrated optoelectronics (Hanewinkel *et al.* 2001).

21.4.2 Ge photodiodes

Ge is categorized into indirect gap material, like Si, with an indirect band gap of 0.64 eV (0.66 eV at RT) and a direct band gap of ~0.8 eV (Fig. 21.4(a)). The difference is fairly small compared with that of silicon (indirect 1.12 eV, direct 3.4 eV); the absorption coefficient is then as large as, and the absorption length is as small as, those of the direct-gap III–V materials. Moreover, a relatively small band gap makes germanium a possible photodiode material up to $\lambda = 1.6$ μm. However, Ge photodiodes on an Si wafer have relatively large dark current, which is the main noise source when no light is shed on a photodiode. This results from threading dislocations in the Ge layer at the interface with Si due to the large lattice mismatch of ~4% between Ge and Si. Those dislocations are believed to require high temperatures of 800–900°C to anneal out (Colace *et al.* 1998). With low temperature oxide Ge p-i-n photodiode can be fabricated on the Si platform (Ishikawa *et al.* 2005, Liu



21.4 Band gap structures of (a) Ge, and (b) strained n-type Ge.

et al. 2005). Further effort is required to reduce the preparing temperature of Ge on Si below ~450°C for the CMOS back-end process.

Ahn *et al*. (2007) reported that waveguide-coupled Ge p-i-n photodiodes have high efficiency (~90%) over a wide range of wavelengths beyond the direct band gap of Ge, resulting in a responsivity of 1.08 *A*/*W* for 1550-nm light. A surface-plasmon antenna with a C-shaped nano-aperture and a bulk Ge substrate demonstrated enhanced photocurrent that was 2–5 times that of a square aperture of the same area (Tang *et al*. 2006). A Ge photodiode with a smaller dipole antenna was also examined (Tang *et al*. 2008). The capacitance of these photodiodes could be less than 10 aF (Latif *et al*. 2009). An avalanche mechanism is occasionally employed to enhance photocurrent by applying a higher reverse voltage (e.g., ~25 V) to increase the carrier concentration due to impact ionization in a semiconductor. Assefa *et al*. (2010) reported that a Ge photodetector using an avalanche mechanism was operated at 1.5 V by strongly generating a non-uniform electric field. The region of impact ionization was reduced to 30 nm and the amplification noise was reduced by over 70%.

21.4.3 Infrared sensors

The lattice mismatch generates the tensile strain in the Ge layers, which reduces the band gap of Ge. A calculation shows that the indirect–direct transition of Ge is brought about when the biaxial strain is 1.4% (Ishikawa and Wada 2010). The band gap at the transition is 0.59 eV which corresponds to a wavelength longer than 2 μm.

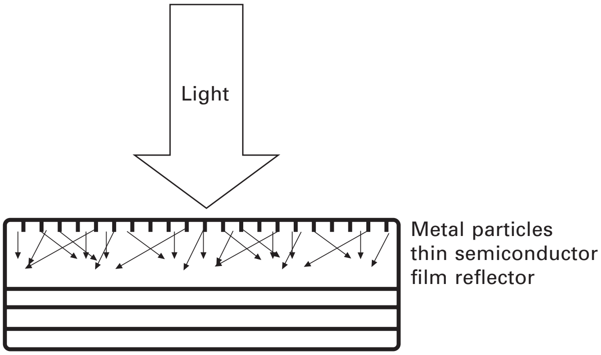
For longer-wavelength applications, such as far-infrared to sub-millimeter wave detection, bolometers employ the sensitivity of the electric resistivity of the temperature change by photon absorption has been widely used (Donati 1999, Rogalski 2001). Amorphous Si (Syllaios *et al*. 2000) and amorphous SiGe with lower resistance (García *et al*. 2004) are used as cost-effective thermistor materials. Series polycrystalline Si p-n junction diodes have also been applied to an infrared imaging device (Ishikawa *et al*. 1999).

21.4.4 Solar cells

Light confinement also increases the effective absorption in silicon solar cells. Unlike photodiodes, solar cells do not require low capacitance for high-speed responses but they require large areas and broad-wavelength response. An important requirement, especially for the silicon solar cell, is that the semiconductor volume should be small, which would cut material costs and total energy consumption, including that used in the process to produce silicon from silica. This makes the light in solar cells confined along the direction of thickness (normal to the surface). A spherical solar cell using a Si sphere with a diameter of 1 mm has also been developed to introduce a low-waste and low-cost fabrication process without cutting and polishing (Minemoto *et al*. 2005).

To make the Si layer thinner to achieve low-cost solar cells, submicron-sized texturing of the substrate (McCann *et al*. 2001, Widenborg and Aberle 2007) has been reported to increase efficiency with polycrystalline Si that is a few micrometers thick. Yablonovitch and Cody (1982) pointed out that the geometry of the silicon surface can be simply roughened or precisely etched to various shapes. Plasmonics (Atwater and Polman 2010) may further reduce the thickness of silicon (Fig. 21.5). Silver (Ag) and gold (Au) nano-particles (Hagglund *et al*. 2008, Derkacs *et al*. 2006) and elliptical Au (Pillai *et al*. 2007) for enhancing surface plasmons have also been reported. Both metal (Au) and dielectric (SiO_2) nano-particles are found to improve the photovoltaic performance (Matheu *et al*. 2008). Silicon nanowire-based solar cells have been developed with the expectation of comparable efficiency to bulk crystalline Si (Tsakalakos *et al*. 2007). Tian *et al*. (2007) developed coaxial silicon nanowire p-i-n type solar cells and succeeded to drive functional nano sensors and logic gates. Li *et al*. (2009) found the addition of a Si nano-pillar array to Si thin film enhanced the light absorption.

Hydrogenated amorphous silicon (a-Si:H) has optical characteristics resembling those expected of a crystal with a direct band gap of 1.5–1.8 eV depending on the degree of disorder in the lattice, whereas hydrogenated microcrystalline silicon (μc-Si:H) has an indirect band gap of ~1.1 eV (Cody *et al*. 1981). Furthermore, it offers a larger absorption coefficient because selection rules that apply to optical transition in a crystal are related in amorphous materials. The micromorph tandem solar cell concept consists of an a-Si:H top cell and a μc-Si:H bottom cell and has been developed to absorb both the visible and infrared parts of light efficiently (Cody *et al*. 1981, Meier *et al*. 2004). Hydrogenated amorphous silicon–germanium (a-SiGe:H)



21.5 Structure of plasmon solar cell.

has also been developed for this purpose (Deng *et al.* 2000). Hybrid solar cells with an a-Si:H film on a single crystal Si have been developed and commercialized (Tanaka *et al.* 2003, Yamamoto *et al.* 2004).

21.5 Light sources

Several cost issues have emerged with III–V compound semiconductors suitable for photonic devices with Si electronics. Some of these have originated from dislocation imperfections due to large lattice mismatches and delamination due to the difference in thermal expansion coefficients, and also the diffusion of Si in III–V, when III–V crystal is epitaxially grown on an Si substrate. The saturated vapor pressure and temperature for crystal growth of compound semiconductors differ vastly from that of Si. The insertion of a graded buffer layer of Ge has been investigated to relax the strain and stress due to lattice mismatch, although this process is not yet standard in Si CMOS fabrication. Once an on-chip light source that consists of Group IV elements such as Si and Ge is achieved for practical use, we have the possibility of developing relatively new applications for ultra-small opto-electronic conversion devices.

21.5.1 Si light sources

Since Si is an indirect-band gap semiconductor, a phonon-assist mechanism is required to recombine the exited electrons with holes. The probability of optical transitions assisted by phonons is low and there are too many challenges to increasing the efficiency of light emission for it to be practical. Therefore, it had long been believed that Si could not have optical gain because of the parasitic absorption process due to free carriers (Dumke 1962).

This belief has recently been criticized theoretically (Trupke *et al.* 2003). Then, Chen *et al.* (2004) reported stimulated emission from a p-n Si diode with SiO_2 particles at its junction. Nano-crystalline, porous silicon, and strained and disordered silicon have aimed at making optical transition possible without the conservation of momentum by introducing imperfections in the crystal to break or lower translational symmetry. Emission intensity often increases when Si is nano-structured or Si is a constituent of chemical compounds. Confined electrons increase the uncertainty of the wave number in such cases, and we can then expect the selection rule for the wave number to be relaxed. This can be explained by the suppression of non-radiative recombination in nano-particles (Shaklee and Nahory 1970). Saito *et al.* (2006) reported a light-emitting field effect transistor with a lateral p-n junction with ultra-thin silicon. A first-principle calculation concluded a positive optical gain from ultrathin Si film (Suwa and Saito 2009). Purcell (1946) pointed out the enhancement of spontaneous emission rates of atoms when they are matched in a resonant cavity. Photonic band gaps, which prevent the existence of photonic modes for certain directions of propagation, are high-quality resonant cavities. Fujita *et al.* (2008) demonstrated that a photonic crystal nano-cavity enhanced light emission from Si. Silicon slot waveguides are also expected to enhance the light emission (Jun *et al.* 2009).

The emission of infrared light from mixed-phase materials such as strained $Si_{1-y}Ge_y$ film on an $Si_{1-x}Ge_x$ substrate (where $y < x$) (Usami *et al.* 1995) and silicide such as $FeSi_2$ (Maeda 2007) have also been reported. Liu *et al.* (2007) reported that the induced tensile stress in epitaxial Ge on Si shrinks the indirect bandgap of Ge. Sun *et al.* (2009a) then demonstrated that the photoluminescence intensity of Ge-on-Si increases with n-type doping, which is explained by a higher electron population in the direct Γ-valley (Fig. 21.4(b)). They also demonstrated electroluminescence from Ge-on-Si (Sun *et al.* 2009b). Finally, Liu *et al.* (2010) reported the observation of lasing from the direct gap transition of Ge-on-Si at room temperature. Tensile strain shrinks the direct gap of Ge to 0.76 eV, and the Ge was doped with 1×10^{19} cm^{-3} phosphorus.

21.5.2 Doping rare earth

Doped rare-earth ions as emission centers act as intra-shell transition where the emission of photons is associated with intra-4f transitions of rare earth ions such as Er^{3+}. Although the dipole transition between electronic levels with the same principal quantum number is inhibited with free ions, dipole transition is permitted when the ions are in a solid where the electronic level is modified by the crystal field. Er atoms in Si tend to aggregate. Furthermore, electrons trapped by Er^{3+} receive kinetic energy to jump in the conduction band and are de-excited, thereby giving rise to temperature quenching, which reduces the emission at RT to ~1/1000. The self-assembled fabrication of Er–Si–O has been demonstrated to obtain infrared luminescence (Isshiki *et al.*, 2004). Isshiki *et al.* (2010) reported the up-conversion emission from ErSiO superlattice crystal. Polman *et al.* (2004) demonstrated optically pumped lasing of SiO_2 troidal microcavities doped with Er^{3+}.

21.6 Future trends

Optical communication is believed to offer a better solution than electrical communication using Cu wires or cables in the sense that it yields lower propagation loss. It also yields a higher data rate density (bits/s/cm^2) than other wireless media such as microwaves since it has a smaller wavelength of light and a wider bandwidth that are suitable for wavelength-division multiplexing (WDM). The obstacles preventing such an ideal communication technology from being introduced into electronics are its higher costs and larger footprints for optical devices. Silicon photonics is expected to offer a low-cost means to mass-produce optical devices by drawing on strengths obtained from the semiconductor industry. Fortunately for optics specialists, the refractive index of Si is as high as ~3.5, and therefore Si or even SiON (n = 1.55–2.0) waveguides can strongly confine light. The propagation loss may be large for such waveguides with a high relative refractive index, $\Delta = (n_{\text{core}} - n_{\text{cladding}})/n_{\text{core}}$, because of strong scattering by the roughness of the interface between the core and cladding. Fortunately again, roughness has been decreased by using advanced Si CMOS processes.

Si photonics is blazing the way for optical sensors and solar batteries to merge with Si CMOS device technology (Kimerling 2005, 2008). To obtain even smaller optical devices, photonic crystals and near-field technologies, including those involving surface-plasmon antennas, are also being introduced to Si photonics. Sub-wavelength optics in Si photonics is providing strong tools to meet the challenges of these emerging applications. Soref (2008) discusses Group IV photonics for the long-wavelength region from mid-infrared to near-THz.

21.7 Sources of further information and advice

An excellent overview of silicon photonics is given in Reed and Knights (2004), where the basic operating principles, especially on waveguides and light sources, are carefully explained against their historical background. Pavesi and Lockwood (2004) and Reed (2008) have provided topics on more technical aspects, including the application of silicon photonics to optical interconnections. Dumon *et al*. (2009) state a foundry approach for silicon photonics. Device requirements for optical interconnects on Si LSIs have been extensively discussed by Miller (2000, 2009). Pavesi (2008) gives a thorough review on light sources for integrated silicon circuits from the basics on lasing to various Si-based materials. More thorough descriptions of light detection are given by Donati (1999) and Rieke (2003). Rogalski (2001) presents an overview of infrared systems and detectors. Green (2005) covers all the important concepts in third-generation photovoltaics. Tsakalakos (2008) discusses bulk nanostructures, quantum wells, nanowires, and quantum dots for photovoltaics. Novotony and Hecht (2006) is an excellent textbook on nanooptics, and Markel and George (2001) present a more professional review of the applications of nanooptics. Yariv and Yeh (2007) also include the application of nanophotonics to optical communication devices.

21.8 References

Ahn D, Hong C-Y, Liu J, Giziewicz W, Beals M, Kimerling L C, Michel J, Chen J and Kärtner X (2007), High performance, waveguide integrated Ge photodetectors, *Opt Express*, 15, 3916–3921.

Anderson D B (1965), Application of semiconductor technology to coherent optical transducers and spatial filters, in Tippett J, *Optical and Electro-optical Information Processing*, Cambridge, MA, MIT Press, 221–334.

Assefa S, Xia F and Vlasov Y A (2010), Reinventing germanium avalanche photodetector for nanophotonic on-chip optical interconnects, *Nature*, 464, 80–85.

Atwater H A and Polman A (2010), Plasmonics for improved photovoltaic devices, *Nature Materials*, 9, 205–213.

Baba T (2008), Slow light in photonic crystals, *Nature Photonics*, 2, 465–473.

Bakoglu H B (1990), *Circuits, Interconnections, and Packaging for VLSI*, Reading, MA, Addison-Wesley.

Bristow A D, Rotenberg N and van Driel H M (2007), Two-photon absorption and Kerr coefficient of silicon for 850–2200 nm, *Appl Phys Lett*, 90, 191104.

Brosi J-M, Koos C, Andreani L C, Waldow M, Leuthold J and Freude W (2008), High-speed low-voltage electro-optic modulator with a polymer-infiltrated silicon photonic crystal waveguide, *Opt Express*, 16, 4177–4191.

Cardenas J, Poitras C B, Robinson J T, Preston K, Chen L and Lipson M (2009), Low loss etchless silicon photonic waveguides, *Opt Express*, 17, 4752–4757.

Chen M J, Yen J L, Li J Y, Chang J F, Tsai S C and Tsai C S (2004), Stimulated emission in a nanostructured silicon pn junction diode using current injection, *Appl Phys Lett*, 84, 2163–2165.

Chen H, Luo X and Poon A W (2009), Cavity-enhanced photocurrent generation by 1.55 μm wavelengths linear absorption in a p-i-n diode embedded silicon microring resonator, *Appl Phys Lett*, 95, 171111.

Chu T, Yamada H, Ishida S and Arakawa Y (2005), Compact 1 × N thermo-optic switches based on silicon photonic wire waveguides, *Opt Express*, 13, 10109–10114.

Cocorullo C, Iodice M, Rendina I and Sarro P M (1995), Silicon thermo-optical micro-modulator with 700 kHz 3db bandwidth, *IEEE Photonics Technol Lett*, 7, 363–365.

Cody G D, Tiedje T, Abeles B, Moustakas T D, Brooks B and Goldstein Y (1981), Disorder and the optical absorption edge of hydrogenated amorphous silicon, *J Physique*, C4, 301–304.

Colace L, Masini G, Galluzi F, Assanto A, Capellini G, Di Gaspare L, Palange E and Evangelisti F (1998), Metal–semiconductor–metal near-infrared light detector based on epitaxial Ge/Si, *Appl Phys Lett*, 72, 3175–3177.

Cotter L K, Drabik T J, Dillon R J and Handschy M A (1990), Ferroelectric-liquid-crystal/silicon-integrated-circuit spatial light modulator, *Opt Lett*, 15, 291–293.

Dakss M L, Kuhn L, Heidrich P F and Scott B A (1970), Grating coupler for efficient excitation of optical guided waves in thin films, *Appl Phys Lett*, 16, 523–525.

de Ridder R M, Wörhoff K, Driessen A, Lambeck P V and Albers H (1998), Silicon oxynitride planar waveguiding structures for application in optical communication, *IEEE J Sel Top Quantum Electron*, 4, 930–937.

Deng X, Liao X, Han S, Povolny H and Agarwal P (2000), Amorphous silicon and silicon geranium materials for high-efficiency triple-junction sol cell, *Solar Energy Mater Sol Cells*, 62, 89–95.

Derkacs D, Lim S H, Matheu P, Mar W and Yu E T (2006), Improved performance of amorphous silicon solar cells via scattering from surface plasmon polaritons in nearby metallic nanoparticles, *Appl Phys Lett*, 89, 093103.

Donati S (1999), *Photodetectors: Devices, Circuits and Applications*, Upper Saddle River, NJ, Prentice-Hall.

Dumke W P (1962), Interband transitions and maser action, *Phys Rev*, 127, 1559–1563.

Dumon P, Bogaerts W, Baets R, Fedeli J-M and Fulbert L (2009), Towards foundry approach for silicon photonics: silicon photonics platform ePIXfab, *Electron Lett*, 45, 581–582.

Emsley M K, Dosunmu O and Ünlü M S (2002), Silicon substrates with buried distributed Bragg reflectors for resonant cavity-enhanced optoelectronics, *IEEE J Sel Topics Quantum Electron*, 8, 948–955.

Espinola R L, Tsai M C, Yardley J T and Osgood R M (2003), Fast and low-power thermooptic switch on thin silicon-on-insulator, *IEEE Photon Technol Lett*, 15, 1366–1368.

Friebele E, Askins C G and Gingerich M E (1984), Effect of low dose rate irradiation on doped silica core optical fibers, *Appl Opt*, 23, 4202–4208.

Fujikata J, Ishi T, Makita K, Baba T and Ohashi K (2005), Highly enhanced speed and efficiency of Si nano-photodiode with a surface-plasmon antenna, *SSDM 2005*, E-3-3.

Fujikata J, Nose K, Ushida J, Nishi K, Kinoshita M, Shimizu T, Ueno T, Okamoto D, Gomyo A, Mizuno M, Tsuchizawa T, Watanabe T, Yamada K and Ohashi K (2008), Waveguide-integrated Si nano-photodiode with surface-plasmon antenna and its application to on-chip optical clock distribution, *Appl Phys Exp*, 1, 022001.

Fujita M, Tanaka Y and Noda S (2008), Light emission from silicon in photonic crystal nanocavity, *IEEE J Sel Topics Quantum Electron*, 14, 1090–1097.

Garcia M, Ambrosio R, Torres A and Kosarev A (2004), IR bolometers based on amorphous silicon germanium alloys, *J Non-Crystalline Solids*, 338–340, 744–748.

Gnan M, Thoms S, Macintyre D S, De La Rue R M and Sorel M (2008), Fabrication of low-loss photonic wires in silicon-on-insulator using hydrogen silsesquioxane elcctron-beam resist, *Electron Lett*, 44, 115–116.

Green M A (2005), *Third Generation Photovoltaics: Advanced Solar Energy Conversion*, Berlin, Springer.

Green W M, Rooks M R, Sekaric L and Vlasov Y A (2007), Ultra compact, low RF power, 10 Gb/s silicon Mach–Zehnder modulator, *Opt Express*, 15, 17106–17113.

Guha B, Kyotoku B B C and Lipson M (2010), CMOS-compatible athermal silicon microring resonators, *Opt Express*, 18, 4, 3438–3493.

Hagglund C, Zach M, Petersson G and Kasemo B (2008), Electromagnetic coupling of light into a silicon solar cell by nanodisk plasmons, *Appl Phys Lett*, 92, 053110.

Han M and Wang A (2007), Temperature compensation of optical microresonators using a surface layer with negative thermo-optic coefficient, *Opt Lett*, 32, 1800–1802.

Hanewinkel B, Knorr A, Thomas P and Koch S W (2001), Near-field optics of nanostructured semiconductor materials, in Markel V A and George T F, *Optics of Nanostructured Materials*, New York, John Wiley, 143–200.

Ishi T, Fujikata J, Makita K, Baba T and Ohashi K (2005), Si nano-photodiode with a surface plasmon antenna, *Jpn J Appl Phys*, 44, L364–L366.

Ishikawa Y and Wada K (2010), Germanium for silicon photonics, *Thin Solid Films*, 518, S83–S87.

Ishikawa T, Ueno M, Endo K, Nakaki Y, Hata H, Sone T and Kimata M (1999), Low-cost 320 × 240 uncooled IRFPA using conventional silicon IC process, *Opto-electron Rev*, 7, 297–303.

Ishikawa Y, Wada K, Liu J, Cannon D D, Luan H -C, Michel J and Kimerling L C (2005), Strain-induced enhancement of near-infrared absorption in Ge epitaxial layers grown on Si substrate, *J Appl Phys*, 98, 013501.

Isshiki H, de Dood M J A, Polman A and Kimura T (2004), Self-assembled infrared-luminescent Er–Si–O crystallites on silicon, *Appl Phys Lett*, 85, 4343–4345.

Isshiki H, Ushiyama T and Kimura T (2010), Demonstration of ErSiO superlattice crystal waveguides towards optical amplifiers and emitters, *Phys Stat Sol (a)*, 205, 52–55.

Jacobsen R S, Andersen K N, Borel P I, Fage-Pedersen J, Frandsen L H, Hansen O, Kristensen M, Lavrinenko A V, Moulin G, Ou H, Peucheret C, Zsigri B and Bjarklev A (2006), Strained silicon as a new electro-optic material, *Nature*, 44, 199–202.

John S (1987), Strong localization of photons in certain disordered dielectric superlattices, *Phys Rev Lett*, 58, 2486–2489.

Jongthanmmanurak S, Liu J F, Wada K, Cannon D D, Danielson D T, Pan D, Kimerling L C and Michel J (2006), Large electro-optic effect in tensile stressed Ge-on-Si films, *Appl Phys Lett*, 89, 161115.

Jun Y C, Briggs R M, Atwater H A and Brongersma M L (2009), Broadband enhancement of light emission in silicon slot waveguides, *Opt Express*, 17, 7479–7490.

Kao C K and Davies T W (1968), Spectroscopic studies of ultra low loss optical glasses, *J Sci Instrum*, 1, 1063–1068.

Kimerling L (2005), *Microphotonics: Hardware for the Information Age*, The Microphotonics Center at MIT. Available from http://www.signallake.com/innovation/MicrophotonicsCommRoadmap2005.pdf [Accessed 31 May 2010].

Kimerling L (2008), 'Foreword', in Reed G T, *Silicon Photonics: the State of the Art*, New York, John Wiley & Sons, xi–xv.

Kuo Y-H, Lee Y K, Ge Y, Ren S, Roth J E, Kamins T I, Miller D A B and Harris J S (2005), Strong quantum-confined Stark effect in germanium quantum-well structures on silicon, *Nature*, 437, 1334–1336.

Latif S, Kocabas S E, Tang L, Dabaes C and Miller D A B (2009), Low capacitance CMOS silicon photodetectors for optical clock injection, *Appl Phys A*, published online, Springer-Verlag, 2009.

Lee S Y and Feigelson R S (1999), *c*-axis lithium niobate thin film growth on silicon using solid-source metalorganic chemical vapor deposition, *J Mater Res*, 14, 6, 2662–2667.

Lee S-S, Garner S M, Chuyanov V, Zhang H, Steier W H, Wang F, Dalton L R, Udupa A H and Fetterman H R (2000), Optical intensity modulator based on a novel electrooptic polymer incorporating a high *mb* chromophore, *IEEE J Quantum Electron*, 36, 5, 527–532.

Li J, Yu H Y, Wong S M, Zhang G, Sun X, Lo P G-Q and Kwong D-L (2009), Si nanopillar array optimization on Si thin films for solar energy harvesting, *Appl Phys Lett*, 95, 033102.

Lipson M (2006), Compact electro-optic modulators on a silicon chip, *IEEE J Sel Topics Quantum Electron*, 12, 1520–1526.

Liu J, Cannon D D, Wada K, Ishikawa Y, Jongthammanurak S, Danielson D T, Michel J and Kimerling L C (2005), Tensile strained Ge p-i-n photodetectors on Si platform for C and L band telecommunications, *Appl Phys Lett*, 87, 011110.

Liu J, Sun X, Pan D, Wang X X, Kimerling L C, Koch T L and Michel J (2007), Tensile-strained, n-type Ge as a gain medium for monolithic laser integration on Si, *Opt Express*, 15, 11272.

Liu J, Beals M, Pomerene A, Bernardis S, Sun R, Cheng J, Kimerling L C and Michel J (2008), Ultralow energy, integrated GeSi electroabsorption modulator on SOI, *GFP 2008*, WA4.

Liu J, Sun X, Camacho-Aguilera R, Kimerling L and Michel J (2010), A Ge-on-Si laser operating at room temperature, *Opt Lett*, 35, 679–681.

Lončar M, Doll T, Vučković J and Scherer A (2000), Design and fabrication of silicon photonic crystal optical waveguides, *J Lightwave Technol*, 18, 1402–1411.

Maeda Y (2007), Semiconducting β-$FeSi_2$ towards optoelectronics and photonics, *Thin Solid Films*, 515, 8118–8121.

Magen N, Kolodny A, Weiser U and Shamir N (2004), Interconnect-power dissipation in a microprocessor, *SLIP '04*, 7–13.

Manipatruni S, Xu Q, Schmidt B, Shakya J and Lipson M (2007), High speed carrier injection 18 Gb/s silicon micro-ring electro-optic modulator, *LEOS 2007*, WO2.

Manipatruni S, Dokania R K, Schmidt B, Sherwood-Droz N, Poitras C B, Aspel A B and Lipson M (2008), Wide temperature operation of micrometer-scale silicon electro-optic modulators, *Opt Lett*, 33, 2185–2187.

Markel V A and George T F (2001), *Optics of Nanostructured Materials*, New York, John Wiley & Sons.

Matheu P, Lim S H, Derkacs D, McPheeters and Yu E T (2008), Metal and dielectric nanoparticles scattering for improved optical absorption in photovoltaic devices, *Appl Phys Lett*, 93, 113108.

McCann M J, Catchpole K R, Weber K J and Blackers A W (2001), A review of thin-film crystalline silicon for solar cell applications. Part 1: Native substrates, *Sol Energy Mater Sol Cells*, 68, 135–171.

Meier J, Spitznagel J, Kroll U, Bucher C, Faÿ S, Moriarty T and Shah A (2004), Potential of amorphous and microcrystalline silicon solar cells, *Thin Solid Films*, 451–452, 518–524.

Meindl J D, Davis J A, Zarkesh-Ha P, Patel C S, Martin K P and Kohl P A (2002), Interconnect opportunities for gigascale integration, *IBM J R&D*, 46, 245–263.

Miller D A B (2000), Rationale and challenges for optical interconnects to chips, *Proc IEEE*, 88, 728–749.

Miller D A B (2009), Device requirements for optical interconnects to silicon chips, *Proc IEEE*, 97, 1166–1185.

Minemoto T, Okamoto C, Omae S, Murozono M, Takakura H and Hamakawa Y (2005), Fabrication of spherical silicon solar cells with semi-light-concentration system, *Jpn J Appl Phys*, 44, 4820–4824.

Nakada M, Ohashi K and Akedo J (2004), Electro-optical properties and structures of (Pb, La)(Zr, Ti)O_3 and $PbTiO_3$ films prepared using aerosol deposition method, *Jpn J Appl Phys*, 43, 6543–6548.

Novotony L and Hecht B (2006), *Principles of Nano-optics*, Cambridge, Cambridge University Press.

Ohashi K and Fujikata J (2009), Photodetector using surface-plasmon antenna for optical interconnect, *Mater Res Soc Symp Proc*, 1145-MM01-05.

Ohashi K, Fujikata J, Nakada M, Ishi T, Nishi K, Yamada Y, Fukaishi M, Mizuno M, Nose K, Ogura I, Urino Y and Baba T (2006), Optical interconnect technologies for high-speed VLSI chips using silicon nanophotonics, *Digest of Technical Papers ISSCC 2006*, 426–427.

Ohashi K, Nishi K, Shimizu T, Nakada M, Fujikata J, Ushida J, Torii S, Nose K, Mizuno M, Yukawa H, Kinoshita M, Suzuki N, Gomyo A, Ishi T, Okamoto D, Furue K, Ueno T, Tsuchizawa T, Watanabe T, Yamada K, Itabashi S and Akedo J (2009), On-chip optical interconnect, *Proc IEEE*, 97, 7, 1186–1198.

Ohtaka K (1979), Energy band of photons and low-energy photon diffraction, *Phys Rev B*, 19, 5057–5067.

Pavesi L (2008), Silicon-based light sources for silicon integrated circuits, *Adv Opt Technol*, 2008, 416926.

Pavesi L and Lockwood D J (2004), *Silicon Photonics*, Berlin, Heidelberg, Springer-Verlag.

Pillai S, Catchpole K R, Trupke T and Green M A (2007), Surface plasmon enhanced silicon solar cells, *J Appl Phys*, 109, 093105.

Polman A, Min B, Kalkman J, Kippenberg T J and Vahala (2004), Ultralow-threshold erbium-implanted toroidal microlaser on silicon, 84, 7, 1037–1039.

Purcell E M (1946), Spontaneous emission probabilities at radio frequencies, *Phys Rev*, 69, 681.

Qian Y, Kim S, Song J and Nordin G P (2006), Compact and low loss silicon-on-insulator rib waveguide 90° bend, *Opt Express*, 14, 6020–6028.

Reed G T (2008), *Silicon Photonics: the State of the Art*, New York, John Wiley & Sons.

Reed G T and Knights A P (2004), *Silicon Photonics: an Introduction*, New York, John Wiley & Sons.

Rieke G (2003), *Detection of Light: from the Ultraviolet to the Submillimeter*, Cambridge, Cambridge University Press.

Rogalski A (2001), Infrared detectors at the beginning of the next millennium, *Optoelectron Rev*, 9 173–187.

Roth J E, Fidaner O, Shaevitz R K, Kuo Y-H, Kamins T I, Harris J S and Miller D A B (2007), Optical modulator on silicon employing germanium quantum wells, *Opt Express*, 15, 9, 5851–5859.

Saito S-I, Hisamoto D, Shimizu H, Hamamura H, Tsuchiya R, Matsui Y, Mine T, Arai T, Sugii N, Torii K, Kimura S and Onai T (2006), Electroluminescence from ultrathin silicons, *Jpn J Appl Phys*, 45, L679–L682.

Schaevitz R K, Roth J E, Ren S, Fidaner O and Miller D A B (2008), Material properties of Si-Ge/Ge quantum wells, *IEEE J Sel Topics Quantum Electron*, 14, 4, 1082–1089.

Schaub J D, Li R, Show C L, Campbell J C, Neubeck G W and Danton J (1999), Resonant-cavity-enhanced high-speed Si photodiode grown by epitaxial lateral overgrowth, *IEEE Photon Technol Lett*, 11, 1647–1649.

Shaklee K L and Nahory R E (1970), Valley-orbit splitting of free excitons? The absorption edge of Si, *Phys Rev Lett*, 24, 942–945.

Shimizu T, Nakada M, Tsuda H, Miyazaki H, Akedo J and Ohashi K (2009), Gigahertz-rate optical modulation on Mach–Zehnder PLZT electro-optic modulators formed on silicon substrates by aerosol deposition, *IEICE Electron Exp*, 6, 1669–1675.

Sigalas M M, Ho K-M, Biswas R and Soukoulis C M (2001), Photonic crystals, in Markel V A and George T F, *Optics of Nanostructured Materials*, New York, John Wiley & Sons, 1–38.

Soref R A (2008), Towards silicon-based longwave integrated optoelectronics, *Proc SPIE*, 6898–5.

Soref R A and Bennet B R (1987), Electrooptical effects in silicon, *IEEE J Quantum Electronics*, 23, 123–129.

Soref R A and Lorenzo J P (1985), Single-crystal silicon: a new material for 1.3 and 1.6 μm integrated-optical components, *Electron Lett*, 21, 953–954.

Soref R A, Emelett S J and Buchwald W R (2006), Silicon waveguided components for the long-wave infrared region, *J Opt A: Pure Appl Opt*, 8, 840–848.

Sun X, Liu J, Kimmerling L C and Michel J (2009a), Direct gap photoluminescence of n-type tensile-strained Ge-on-Si, *Appl Phys Lett*, 95, 011911.

Sun X, Liu J, Kimmerling L C and Michel J (2009b), Room temperature direct band gap electroluminescence from Ge-on-Si light emitting diode, *Opt Lett*, 34, 8, 1198–1200.

Suwa Y and Saito S-I (2009), Intrinsic optical gain of ultrathin silicon quantum wells from first-principles calculations, *Phys Rev*, B79, 233308.

Syllaios A J, Schimert T R, Gooch R W, McCardel W L, Ritchey B A and Tregilgas J H (2000), Amorphous silicon microbolometer technology, *Mat Res Soc Symp Proc*, 609, A14.4.1.

Tamir T (1975), Beam and waveguide couplers, in Tamir T, *Integrated Optics*, Berlin, Springer-Verlag, 83–137.

Tanabe T, Sumikura H, Taniyama H, Shinya A and Notomi M (2010), All-silicon sub-Gb/s telecom detector with low dark current and high quantum efficiency, *Appl Phys Lett*, 96, 101103.

Tanaka M, Okamoto S, Tsuge S and Kiyama S (2003), Development of HIT solar cells with more than 21% conversion efficiency and commercialization of highest performance HIT modules, *Third World Conf on Photovoltaic Energy Conversion*, 955.

Tang L, Miller D A B, Okyay A K, Matteo J A, Yuen Y, Saraswat K C and Hesselink L (2006), C-shaped nanoaperture-enhanced germanium photodetector, *Opt Lett*, 31, 1519–1521.

Tang L, Kocabas S E, Latif S, Okyay A K, Ly-Gagnon D-S, Saraswat K C and Miller D A B (2008), Nanometre-scale germanium photodetector enhanced by a near-infrared dipole antenna, *Nature Photonics*, 2, 226–229.

Tian B, Zheng X, Kempa T J, Fang Y, Yu N, Yu G, Huang J and Lieber C M (2007), Coaxial silicon nanowires as solar cells and nanoelectronic power sources, *Nature*, 449, 885–890.

Tokushima M, Yamada H and Arakawa Y (2004), 1.5-μm-wavelength light guiding in waveguides in square-lattice-of-rod photonic crystal slab, *Appl Phys Lett*, 84, 4298–4300.

Tomme E V, Daele P P V, Baets R G and Lagasse P E (1991), Integrated optic devices based on nonlinear optical polymers, *IEEE J Quantum Electronics*, 27, 778–787.

Trupke T, Green M A and Würfel P (2003), Optical gain in materials with indirect transitions, *J Appl Phys*, 93, 9058–9061.

Tsakalakos L (2008), Nanostructures for photovoltaics, *Mater Sci Eng*, R 62, 175–189.

Tsakalakos L, Balch J, Fronheiser J and Korevaar B A (2007), Silicon nanowire solar cells, *Appl Phys Lett*, 91, 233117.

Tsang H K, Wong C S and Liang T K (2002), Optical dispersion, two-photon absorption and self-phase modulation in silicon waveguides at 1.5 μm wavelength, *Appl Phys Lett*, 80, 416–418.

Tsuchizawa T, Yamada K, Fukuda H, Watanabe T, Uchiyama S and Itabashi S (2006), Low-loss Si wire waveguides and their application to thermooptic switches, *Jpn J Appl Phys*, 45, 6658–6662.

Tsuchizawa T, Watanabe T, Yamada K, Fukuda H, Itabashi S, Fujikata J, Gomyo A, Ushida J, Okamoto D, Kenichi N and Ohashi K (2008), Low-loss silicon oxynitride waveguides and branches for the 850-nm-wavelength region, *Jpn J Appl Phys*, 47, 6739–6743.

Usami N, Isshiki F, Nayaku D K, Shiraki Y and Fukatsu S (1995), Enhancement of radiative recombination in Si-based quantum wells with neighboring confinement structures, *Appl Phys Lett*, 67, 524–526.

Valetta S, Renard S, Jadot J P, Gidon P and Erbeia C (1989), Silicon-based integrated optics technologies, *Solid State Technol*, 32, 69–75.

Vlasov Y, Green W M J and Xia F (2008), High-throughput silicon nanophotonic wavelength-insensitive switch for on-chip optical networks, *Nature Photonics*, 2, 242–246.

Widenborg P I and Aberle A G (2007), Polycrystalline silicon thin-film solar cells on AIT-textured glass superstrates, *Advances in OptoElectronics*, 2007, 24584.

Wülbern J H, Petrov A and Eich M (2009), Electro-optical modulator in a polymer-infiltrated silicon slotted photonic crystal waveguide heterostructure resonator, *Opt Express*, 17, 304–313.

Xu Q, Almeida V R, Panepucci R R and Lipson M (2004), Experimental demonstration of guiding and confining light in nanometer-size low-refractive-index material, *Opt Lett*, 29, 1626–1628.

Yablonovitch E (1987), Inhibited spontaneous emission in solid-state physics and electronics, *Phys Rev Lett*, 58, 2059–2062.

Yablonovitch E and Cody G D (1982), Intensity enhancement in textured optical sheets for solar cells, *IEEE Trans Electron Devices*, 29, 300–305.

Yamada H, Chu T, Ishida S and Arakawa Y (2006), Si photonic wire waveguide devices, *IEEE J Sel Topics Quantum Electron*, 12, 1371–1379.

Yamamoto K, Nakajima A, Yoshimi M, Sawada T, Fukuda S, Suezaki T, Ichikawa M, Koi Y, Goto M, Meguro T, Matsuda T, Kondo M, Sakaki T and Tawada Y (2004), A high efficiency thin film silicon solar cell and module, *Solar Energy*, 77, 939–949.

Yariv Y and Yeh P (2007), *Photonics*, New York, Oxford University Press.

Zalewski E F and Duda C R (1983), Silicon photodiode device with 100% external quantum efficiency, *Appl Opt*, 22, 2867–2873.

22

Spintronics of nanostructured manganese germanium (MnGe) dilute magnetic semiconductor

K. L. WANG and F. XIU, University of California, Los Angeles, USA and A. P. JACOB, Intel Corporation, USA

Abstract: With the apparent limit of scaling on CMOS microelectronics fast approaching, spintronics has received enormous attention as it promises next-generation nanometric magnetoelectronic devices; particularly, the electric field control of the ferromagnetic transition in dilute magnetic semiconductor (DMS) systems offers magnetoelectronic devices a potential for low power consumption and low variability. Special attention has been given to technologically important group IV semiconductor based magnetic materials, with a prominent position for MnGe. Since the first claim of the realization of a MnGe DMS, tremendous efforts have been concentrated on enhancing the Curie temperature and on interpreting the observed ferromagnetism in terms of DMS theories. In this chapter, we will first review the current theoretical understanding on ferromagnetism in MnGe DMS, pointing out the possible physics models underlying the complicated ferromagnetic behavior of MnGe. Then we carry out detailed analysis of MnGe thin films grown by molecular beam epitaxy. We show that with zero- and one-dimensional quantum structures, superior magnetic properties of MnGe compared with bulk films can be obtained. More importantly, with MnGe nanostructures, such as quantum dots, we demonstrate a field-controlled ferromagnetism up to 100 K. The controllability of ferromagnetism in this material system presents a major step towards Ge based spintronic devices working at ambient temperatures.

Key words: dilute magnetic semiconductors, spintronics, MnGe, quantum dots, hole-mediated ferromagnetism, electrical field controlled ferromagnetism.

22.1 Introduction

Dilute magnetic semiconductors (DMSs) attract tremendous interest as emerging candidates for the microelectronics industry due to their uniqueness in exhibiting spin-dependent magneto-electro-optical properties.[1,2] Their distinctive material characteristics such as spin-dependent coupling between semiconductor bands and the localized states promises the magnetoelectric effect – the modulation of magnetic properties by an applied electric field – in semiconductors[1,2]. Thus, a wide variety of semiconductor devices can be envisaged,[3] such as spin-polarized light-emitting diodes, lasers[4,5] and spin-transistor logic devices.[2,6,7] The development of these devices is considered as a possible route for extending the semiconductor scaling roadmap to spin-added electronics.[8]

DMS materials can be developed by alloying semiconductors with several percentages of magnetic transition elements (with 3-*d* orbitals), such as Fe, Mn, Co, Ni, and V.[9] The anticipated outcome of this doping is the fact that the transition elements occupying substitutional sites hybridize with the semiconductor host via the *sp-d* exchange interaction and help to enhance the spin-dependent transport, which in turn increases both the magnetization and the Curie temperature (T_c) of the semiconductor system.[10] This means that transition metal doping can generate strong spin-dependent coupling states in semiconductor systems, which may be further modulated by an electric field.[2] Since the bandgap engineering and crystal structures of a DMS material are compatible with other semiconductors, they offer significant integration advantages as well.[3]

The dependence of the transition temperature on the density of mobile carriers is considered as good evidence for a genuine DMS. However, the key challenge for the growth of DMS materials is the ability to control synthesis of transition metals in semiconductors, aimed at increasing the doping concentration in order to realize room-temperature ferromagnetism. Thus, achieving both semiconducting and ferromagnetic properties at room temperature has become a great challenge. Unfortunately, this is hampered by the low solubility of transition metal species with a strong tendency to form metallic clusters or heterogeneous regions within the matrix. So the primary research on DMS growth is to produce samples with uniform doping and free of metallic precipitates.

Substantial theoretical and experimental work has been carried out in the past two decades, either to predict or to measure the magnetic properties of a variety of semiconductor materials, including binary and ternary compounds such as III–V, II–VI, IV–VI, VI oxides, III nitrides and group IV systems. Among them, (Ga, Mn)As became the first and the most studied DMSs since 1996;[11] but to our knowledge, some studies lacked the threshold doping profile to push the Curie temperature up to room temperature and above (it is currently around 190 K[12–15]). Since then, substantial research has been carried out in III–V and II–VI semiconductors to improve the T_c by increasing the solubility limit and meanwhile minimizing the self-compensation effect of magnetic structures, using radical techniques such as delta doping of modulation-doped quantum structures.[16–19] Another approach to achieving high-T_c DMS is to use transition metal-doped wide bandgap nitrides and oxides, such as GaN, AlN, ZnO and TiO_2. In these materials, high magnetic impurity concentrations are needed due to the lack of long-range interaction among magnetic spins. However, there have been only limited reports on the carrier-mediated effect in wide bandgap materials at room temperature.[20–22] Although a considerable amount of work has been done on III–V and II–VI DMSs,[2,23] the lack of high T_c and the difficulty in achieving electric field-modulated ferromagnetism (carrier-mediated ferromagnetism) at or above room temperature deter them from being considered for semiconductor integration.

Group IV semiconductors are of particular interest to spintronics technology because of their enhanced spin lifetime and coherent length due to their low spin-orbit coupling and lattice inversion symmetry.[24,25] There is also considerable interest in transition metal-doped group IV semiconductors[26–36] in the semiconductor industry, owing in part to their excellent compatibility with silicon process technology. Several transition metals – including single elements such as Mn, Cr, Co and Fe, and co-doping with Co–Mn and Fe-Mn – have been used as magnetic dopants in both Si and Ge, either through blanket implantation or through *in-situ* doping during the epitaxial growth process. Both Si and Ge are reported to be ferromagnetic.[33,37,38] The reported experimental results also clearly showed the presence of carrier-mediated ferromagnetism in DMS Ge.[39,40] A recent study predicted that the T_c could be increased by enhancing the substitutional doping of Mn in Ge and Si, via co-doping, by adding conventional electronic dopants such as As or P during the Mn doping process.[41] Theories[31] and experiments[29,42] further suggest that the ferromagnetic transition temperature is related to the ratio of interstitial to substitutional Mn, similar to that in the group III–system.[10]

Considering the fact that nanosystems such as nanowires and quantum dots (QDs) are the versatile building blocks for both fundamental studies in nanoscale and the assembly of present-day functional devices, low-dimensional Ge DMS systems such as MnGe nanowires[43–46] and QDs[47] have also been reported to have room-temperature ferromagnetism. The above results are supported by modeling results,[48] although carrier-mediated exchange has yet to be experimentally verified in nanowires. Similarly, high Curie temperatures have been observed in other transition metal-doped semiconductor nanostructures such as II–VI (including ZnO,[49–51] ZnS[52,53] and CdS[53]) and III–V (including GaN[53,54] and GaAs[55]), but only inconclusive reports have been published on the magneto-electric transport in these nanostructures. This could be attributed to the complicated ferromagnetic properties caused by precipitates, second-phase alloys and nanometer-scale clusters. Therefore, it is important and indispensable to differentiate their contributions. Unfortunately, there is still a lack of fundamental understanding of the growth process and spin-dependent states in these systems that gives an impetus to extensive investigations into these and related material systems.

In this chapter, we will first introduce the state-of-the-art theoretical understanding of ferromagnetism in group IV DMSs, particularly pointing out the possible physics models underlying the complicated ferromagnetic behavior of MnGe. Then, we primarily focus on the magnetic characterizations of epitaxially grown MnGe thin films ranging from three-dimensional thin films to zero-dimensional QDs. It is found that when the material dimension decreases (to zero), it evidences a change of material structure, being free of precipitates, and a significant increase of T_c over 400 K. The preliminary explanation for these behaviors can be attributed to the carrier confinement in the DMS nanostructures which strengthens hole localization and subsequently enhances the thermal stability of magnetic polarons, thus giving rise to a higher T_c than those of bulk films.

Another important aspect of DMS research lies in electric field-controlled ferromagnetism. We emphasize the existing experimental efforts and achievement in this direction, and provide a detailed description of field-controlled ferromagnetism of MnGe nanostructures. We show that the ferromagnetism of MnGe nanostructures, such as QDs, can be manipulated via control of the gate voltage in a gate metal–oxide–semiconductor (MOS) structure (up to 100 K). These experimental data suggest hole-mediated ferromagnetism controlled by a gate bias and promise the potential of using MnGe nanostructures to achieve spintronic devices. Finally, we offer general comments on the development of MnGe DMS systems and possible research approaches to achieve room-temperature field-controlled ferromagnetism for spintronic devices working in ambient environments.

22.2 Theories of ferromagnetism in group IV dilute magnetic semiconductor (DMS)

As in other popular dilute magnetic semiconductors such as (Ga, Mn)As, the formation of magnetic order and the origin of ferromagnetism in MnGe can be described (within the parametric limits) using either complementary mean field approximation models such as Zener–kinetic exchange,[56] the Ruderman–Kittel–Kasuya–Yosida (RKKY)[57] interaction and percolation theory,[58] or density functional theory such as *ab initio* full-potential linearized augmented plane wave (FLAPW)[59] electronic calculations. Interestingly, the Zener–kinetic exchange model proposed in 1950 could be used for interpreting many of the experimental results in transition metal-doped semiconductors[60–62] besides the transition metals themselves. All the above theoretical approaches have tried to address and have to a certain extent successfully answered chemical, electronic and magnetic properties of the transition metal-doped group IV semiconductors within the context of (1) the nature of transition metal impurities, (2) the origin of ferromagnetism, (3) the influence of dopants on the T_c, (4) the solubility limits of carrier density on the chemical nature, (5) the influence of carrier density on the electronic structure, and (6) the influence of different hosts on the electronic structure. Within the above theoretical approaches, these models assume that the ferromagnetism in transition metal-doped Ge systems is mediated through exchange coupling between the carriers and the interacting lattice.

This means that the interaction is through *p-d* orbital hybridization between the *d* levels and the valence band *p* states of Ge. According to these models, as the number of dopant carriers increases both the electronic mobility and the ferromagnetic ordering of the system also increase along with the T_c. Thus, the assumption is that the strong exchange coupling between the *p* and *d* orbitals leads to strong kinetic exchange coupling between the spin-polarized holes and spin-polarized TM substitutional vacancies. Among all the 3-*d* TM impurities (TM = V, Cr, Mn, Fe, Co, Ni), Mn is the most favored dopant. This is because, compared to other dopants, Mn favors less clustering, low non-uniform doping distribution and higher substitutional doping over interstitials. The *ab initio* electronic structure calculations within density functional theory show that such doping behaviors could result in relatively higher carrier hole concentrations and increased magnetic moments.[27,59] Complementary theoretical work using the frozen-magnon scheme also shows that the ferromagnetism is produced only through holes and Mn is a good source for generating holes in Ge systems.[63]

In addition, some of the other factors that determine the ferromagnetic strengths in Ge are (1) disorder of Mn site locations, (2) distance between Mn–Mn atoms, (3) solid solubility limits and preferential surface orientation for interstitial and substitution site formation,[64–66] and (4) co-doping of Co or Cr with Mn (they were found to reduce cluster formation but not to significantly contribute to the ferromagnetism in the system[67,68]). According to full-potential linearized augmented plane wave (FLAPW) calculations made on the RKKY model, the coupling between the Mn atoms could be ferromagnetic (FM) or anti-ferromagnetic (AFM) in nature, but the interaction is very localized to Mn sites and depends on the distance between Mn atoms: the density of states decreases rapidly as the Ge atom moves away from Mn.[59] However, disorder of Mn site locations is found to influence the ferromagnetic strength and the transition temperature. In addition, the impact of inherent disorder is also confirmed both by experiments and by the mean field approximation.[58,65,69]

Some of the above models found it hard to explain the long-range exchange interaction and coupling strength when there is varying dopant concentration and disorder ranging from a high-resistive state to a half-metallic state. On the other hand, there are attempts to find solutions for disorder-dependent magnetic variation using percolation theory-based mean field approximation calculations.[58,70] Percolation theory provides possible explanations for the experimentally observed low values of saturation magnetization in low carrier density MnGe, because the infinite cluster of percolating bound magnetic polarons triggering the long-range ferromagnetic order leaves out a large number of Mn moments at temperatures less than T_c. However, the role of disorder with respect to insulating and conductive metallic phases and their effects on mobility is still not well understood. Conversely, the FLAPW calculations using density functional theory show that Mn favors long-range ferromagnetic alignment over short-range anti-ferromagnetic alignment and this increases with the increase of Mn content, which is in agreement with many of the experimental results.[59] The estimated Curie temperature is within the range of 134–400 K and is in agreement with experimental and theoretical results.[48]

However, there are also conflicting reports that long-range FM order could only happen at low temperature (12 K) and that at low temperature MnGe could show only spin glass behavior due to the inter-cluster formation between the FM Mn-rich clusters.[31,71] With these reports, it is understood that the magnetic phases are more favored over paramagnetic phases by an energy difference of 200 meV/Mn atom.[59] The report also states that the MnGe system is very close to half-metallicity irrespective of the amount of doping concentration.[59,72] Electronic band structure calculations using density of states (DOS) show that the binding energies of the Ge 4*s* states are basically unaffected by the exchange splitting of Mn *d* states and the valence band maximum of the minority spin channel in MnGe reaches E_F at Gamma (Γ) such that the spin gap becomes zero.[59,72] In other words, the DOS shows a valley around E_F and is strictly zero at E_F in the minority spin channel. On the other hand, the energy gap is indirect and the bands around E_F in the majority spin channel arise from Mn *d* and Ge *p* hybridizations so as to give rise to hole pockets closer to the gamma point with a localized magnetic moment of about $3\mu_B$. This could also be explained using chemical and structural configurations of Mn in Ge. The different chemical species determine the relative positions of the anion and cation atomic energy levels in the MnGe while the different anion size dictates essentially the equilibrium lattice constants.[59] This essentially explains the carrier-mediated effect in MnGe even though the system is closer to half-metallic.

The aim of all this theoretical work is to build a foundation for achieving carrier-mediated ferromagnetism at temperatures above room temperature. It is understood that this can be achieved only by growing cluster-free and uniformly doped MnGe samples. Experimental reports show that T_c as high as 400 K is obtained for cluster-free MnGe samples grown using (1) the sub-surfactant epitaxial method,[73] (2) epitaxial QDs,[40] and (3) nanowire structures.[30,43,45] However, there are only very few theoretical studies depicting the relation between the quantum confinement and its influence on the origin of ferromagnetism in low-dimensional DMS structures.[74–76] There is apparently a lack of systematic theoretical studies, particularly on MnGe QDs. Therefore, it will be interesting to investigate the ability of the quantum confinement phenomenon to retain spin polarization in MnGe QDs structures. Studies on III–V QDs using local spin-density approximation (LSDA) show that, unlike in bulk structures, adding a single carrier or multiple carriers in a magnetic QD can strongly change both the total carrier spin and the temperature of the onset of magnetization, and this can be modulated by modifying the quantum confinement and the strength of Coulomb interactions. These theoretical results must have ramification in the high Curie temperature MnGe QDs as well, since our studies also reveal similar characteristics. But not all of the effects, including high T_c, that we observe can be explained without invoking tiny defects either at the interface or on the surface, although conventional transmission electron microscopy (TEM) cannot reveal their existence due to its limited resolution. All of the above material and electronic properties have been studied in order to understand how feasible a DMS material is for carrier-mediated ferromagnetism.

22.3 Growth and characterizations of group IV dilute magnetic semiconductor (DMS) and nanostructures

22.3.1 Overview of Mn doped Ge

As mentioned above, Ge-based DMSs have attracted extensive attention due to the possibility of integrating them with mainstream Si microelectronics, in which they may be used to enhance the functionality of Si integrated circuits.[27] In particular, the hole-mediated effect discovered in MnGe DMS opens up tremendous possibilities to realize spintronic devices, with advantages in reducing power dissipation and increasing new functionalities. To date, there have been many reports on MnGe growth and characterization by molecular beam epitaxy (MBE),[27,28,31,33,42,71,77–84] ion implantation,[85–92] and bulk crystal growth.[37,93] We can divide them here, for easier orientation, into two groups: those that cover fundamental studies of phase formation, ferromagnetism, and transport properties,[26,27,29,31,33,35–37,42,45,46,65,66,71,77–84,93–113] and those that aim to minimize cluster formation and enhance Curie temperatures via co-doping methods.[28,32,34,114–116]

Experimental results show that the Mn doping process in MnGe is complex. Both the T_c and the saturation magnetization depend on the interplay of a variety of factors, which are ultimately determined by growth conditions and the post-annealing process.[33] The concentration and distribution of Mn dopants, the carrier density, and the presence of common defects such as Mn interstitials and Mn clusters significantly influence the magnitude and interactions of the magnetic coupling.[93] In early 2002, Park *et al.*[27] were pioneers in demonstrating ferromagnetic MnGe thin films grown by low-temperature MBE. It was found that the MnGe films were *p*-type semiconductors in character with a hole concentration of 10^{19}–10^{20} cm^{-3}, and exhibited a pronounced extraordinary Hall effect. The T_c of the MnGe thin films increased linearly with Mn concentration from 25 to 116 K (Mn $\leq 3.5\%$). Field-controlled ferromagnetism was also observed in a simple gated structure through application of a gate voltage (±0.5 V), showing a clear hole-mediated effect at 50 K. Such a small gate voltage presents excellent comparability with conventional low-voltage circuitry. The origin of the ferromagnetic order was understood in the frame of local-spin density approximation (LSDA), where strong hybridization between Mn *d* states of T_2 symmetry with Ge *p* states leads to the configuration $e(\uparrow)^2T_2(\uparrow)^2T_2(\downarrow)^1$, with a magnetic moment of $3\mu_B$. However, the LSDA calculations overestimated T_c, which was attributed to the incomplete activation of Mn in experiments and to the absence of hole compensation in the simulation.[27]

Since then, various preparation techniques have been employed to produce Mn-doped Ge DMS, including the aforementioned MBE[28,31,33,42,71,77–84], single-crystal growth,[37,93] and ion implantation,[85–92] aiming to further increase T_c and to obtain electric field controllability at room temperature.[93] Among these efforts, Cho *et al.*[37] reported the synthesis of Mn-doped bulk Ge single crystals with 6% Mn and a high ferromagnetic order at about 285 K via a vertical gradient solidification method. The origin of the high T_c was found to be complex because of the presence of dilute and dense Mn-doped regions. Jaeger *et al.*,[71] however, pointed out that the magnetic properties of these samples were clearly dominated by the presence of the intermetallic compound $Mn_{11}Ge_8$. This was indeed observed by Biegger *et al.*,[93] where similar MnGe crystals were produced via Bridgman's crystal growth technique and intermetallic compounds were found in both Mn-rich and Mn-poor regions. These experiments suggest that the bulk MnGe crystals produced at over 1000°C may not be an appropriate method for MnGe DMS preparation.[93]

Apart from the bulk crystal growth, MBE has been widely recognized as a power tool for DMS growth since it can provide non-equilibrium conditions to enhance the incorporation of Mn dopant.[27] Via this technique, Mn doping was found to be extremely sensitive to growth conditions, particularly to the growth temperature. Excellent work in this respect can be found in references 32, 33, 42, 78, 80, 82 and 84. In general, at low growth temperatures (T_b < 120°C), the diffusion of Mn atoms leads to the formation of Mn-rich nanostructures, such as nanocolumns[81] and nanodots[80] with irregular shapes. These nanostructures could contain a high Mn concentration, up to 38%[33], while the surrounding matrix retains a low Mn concentration, less than 1%. Perfect lattice coherence was found in the proximity of the nanostructures, indicating strong compression for the coherent nanostructure and large tension for the surrounding matrix.[81] These observations are in good agreement with theoretical predictions of spinodal decomposition,[117] which occurs under a layer-by-layer growth mode with strong pairing attraction between Mn atoms and a tendency to surface diffusion of Mn atoms. These theoretical calculations successfully explained the formation of nanostructures in (Ga, Mn)N and (Zn, Cr)Te, [81,117] which can also be readily applied to the MnGe system.[81] The magnetic properties, however, exhibit various characteristics depending on the nanostructure size and the Mn concentration. In most cases, zero-field cooled (ZFC) and field cooled (FC) magnetizations show superferromagnetic properties with a blocking temperature of ~15 K. It is now clear that such a low blocking temperature originates from the coherent nanostructures with a diameter below several nanometers. Curie temperatures of these nanostructures were also identified to be below 170 K and varied with different Mn concentration. No Mn_5Ge_3 metallic phases were observed in the low-temperature regime (T_b < 120°C).

Bougeard *et al.*[80] also found that when the dimension of the nanostructures was further reduced, the films showed no overall spontaneous magnetization down to 2 K. TEM results were interpreted in terms of an assembly of superparamagnetic moments developing in the dense distribution of nanometer-sized clusters. In addition to the single crystalline MnGe, homogeneous amorphous nanoclusters were also observed under certain growth conditions and contributed to ferromagnetic order below 100 K.[81,105] To render a general understanding of MnGe magnetic properties, Jaeger *et al.*[71] presented a detailed study on $Mn_{0.04}Ge_{0.96}$ and $Mn_{0.20}Ge_{0.80}$ thin films grown at low temperatures and related their magnetic behavior to that of spin-glass. A frozen magnetic state at low temperatures was observed and attributed to the formation of nanostructures. This is in good agreement with Devillers *et al.*[81] Therefore, it can be concluded that the MnGe magnetic semiconductor is not a conventional ferromagnet since its magnetic properties are significantly complicated by the formation of nanostructures and metallic clusters.

When the growth temperature falls into the range 120–145°C, the Mn_5Ge_3 nanoclusters start to develop and dominate the magnetic property with a Curie temperature of 296 K.[33,81] This phase, frequently observed at high-temperature growth, gives the most stable (Mn, Ge) alloy. The other stable compound, $Mn_{11}Ge_8$, was also observed in nanocrystallites surrounded by pure Ge.[82] Although they are ferromagnetic, their metallic character considerably jeopardizes their potential for spintronic applications. Under a narrow growth window around 130°C, another phase, $MnGe_2$ nanocolumns, could be developed with ~33% Mn and a T_c beyond 400 K.[33] A remarkable feature of these clusters is the large magneto-resistance due to geometrical effects. Further increasing the growth temperature leads to the dominating magnetic behavior from the Mn_5Ge_3 clusters. As the coherent nanostructures grow larger at high-temperature growth, higher blocking temperatures (~30 K) were also obtained.[81]

To minimize the phase separation, Mn co-doping with Co was attempted to stabilize structures at high Mn doping concentration.[28,34,114] This showed that co-doping with Co can dramatically reduce phase separation and diffusion of Mn within the Ge lattice while it magnetically complements Mn. The measured strain states indicate the critical role played by substitutional Co with its strong tendency to dimerize with interstitial Mn.[34] The highest T_c achieved so far is about 270 K.[28] Similarly, Gareev *et al.*[116] reported on the co-doping of Fe to achieve a T_c of 209 K. Recent polarized neutron reflectivity measurements provided evidence to show no segregation and lower clustering tendencies for higher Fe doping,[115] which further supported the co-doping approach in MnGe.

Ion implantation was also attempted due to its efficiency of Mn incorporation.[91] However, extensive experiments revealed that Mn-rich precipitates were developed and buried in a crystalline Ge matrix[85–92] in a very similar manner to that of the thin films grown by MBE.[32,42,80,82]

Although there has been much progress in producing high-quality MnGe thin films, the formation of uncontrollable metallic clusters yet remains a challenge. High T_c DMS thin films without any precipitation seem to be a critical obstacle for the fabrication of practical spintronic devices functioning at room temperature, not to mention the room-temperature controlled ferromagnetism. Table 22.1 provides a brief summary of the current status of MnGe including the growth method, the T_c, and the field controllability of ferromagnetism.

22.3.2 MnGe by molecular beam epitaxy

To gain further insight into the complex growth of this material system, we have performed detailed structural and magnetic characterizations of MnGe grown by MBE with different thickness. Our experiments showed that the Mn doping behavior and magnetic properties are dramatically different when the material dimension becomes smaller. Bulk MnGe films always tend to form clusters no matter what the growth conditions are. When it comes to one and zero dimensions, however, the cluster formation may possibly be limited because the Mn concentration cannot accumulate enough and the strain can be easily accommodated to minimize the metallic clusters.[83,92]

Table 22.1 A brief summary of MnGe growth and characterization

Mn-doped Ge	Growth method	Mn (%)	T_C (K)	Origin of FM	Field-controlled FM	References
Thin films	MBE	0.6–3.5	25–116	DMS	50 K	27
Bulk single crystal	Sintering	6.2	150–285	Mn_5Ge_3, $Mn_{11}Ge_8$	No	37, 93
Thin films	MBE	<6	~300	Mn_5Ge_3	No	36, 71, 97
Thin films (with $MnGe_2$)	MBE	6	>400	$MnGe_2$ clusters	No	33
Thin films	Implantation	A few	~300	Mn_5Ge_3, $Mn_{11}Ge_8$	No	39, 87, 88, 91

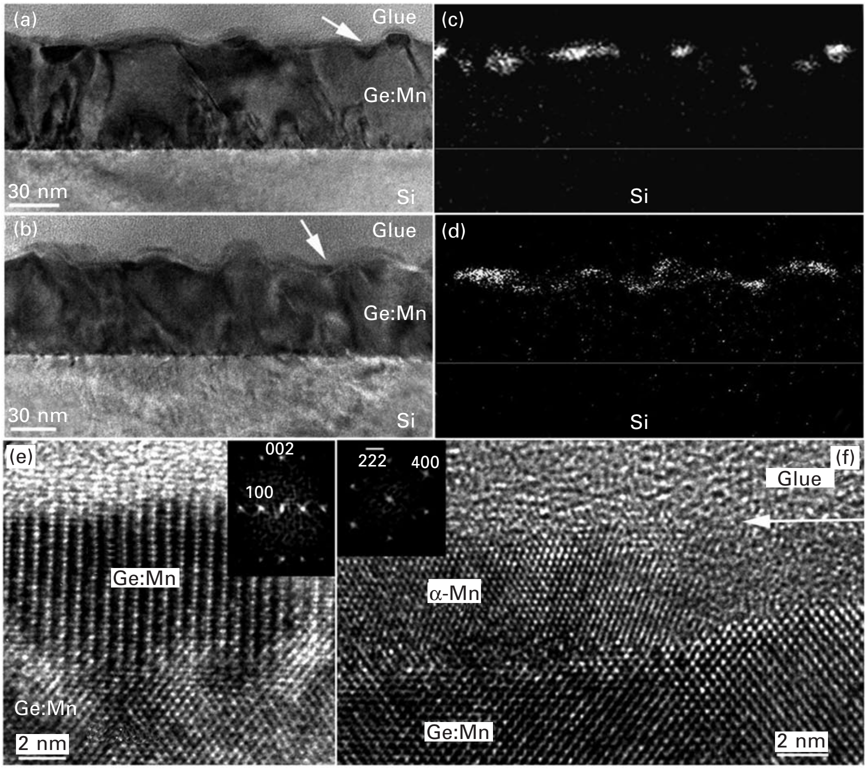
We further show that self-assembled MnGe DMS QDs can be successfully grown on Si substrate with a high T_c in excess of 400 K. Finally, electrical field-controlled ferromagnetism was also demonstrated in MOS structures using these QDs as the channel layers, as elaborated in Section 22.4.

MnGe thin and thick films

Mn-doped Ge films were grown on Si substrates by Perkin–Elmer solid-source MBE. Si substrates were cleaned by H_2SO_4:H_2O_2 (5:3) and 10% HF with a final step of HF etching. The native oxide was removed by a 10 min flash at 800°C in a vacuum chamber. After that, direct growth of 80 nm thick MnGe on Si was obtained at 250°C. A 5 nm Ge capping layer was deposited on the surface in order to prevent MnGe from oxidizing. To probe the annealing effect, identical MnGe thin films were grown repeatedly and followed by an *in-situ* annealing process at 400°C for 30 min. All as-grown and annealed samples were characterized by high-resolution TEM (HR-TEM, FEI Tecnai F30 operating at 200 keV) and a superconducting quantum interference device (SQUID, Quantum Design MPMS XL 5). Cross-sectional HR-TEM specimens along the $\langle 110 \rangle$ plane were prepared by using a tripod technique, followed by a final thinning using a Gatan precision ion polishing system.

Figures 22.1(a) and 22.1(b) show cross-sectional TEM images of the as-grown and annealed 80 nm-thick $Mn_{0.04}Ge_{0.96}$. Non-uniform and discontinued rough films were observed in both cases. Figures 22.1(c) and 22.1(d) are their corresponding Mn elemental maps, respectively. Mn is seen to concentrate on top for both samples. Mn-containing clusters are confirmed near the top of the crystalline layer as indicated by the white arrow in Fig. 22.1(a). A HR-TEM image clearly indicates that a Mn-containing cluster is formed on top of the Ge lattice as shown in Fig. 22.1(e). Using the lattice constant of Ge as the reference, the lattice spacings of the Mn-containing cluster can be determined as 0.62 and 0.25 nm, which are matched with the distances of (100) and (002) atomic planes of the hexagonal Mn_5Ge_3 phase. The fast Fourier transformation (FFT) map shown in the inset of Fig. 22.1(e) also indicates that the cluster belongs to the hexagonal Mn_5Ge_3 phase. Mn_5Ge_3 clusters may be formed when the Mn concentration is sufficiently high for the nucleation of clusters.[83]

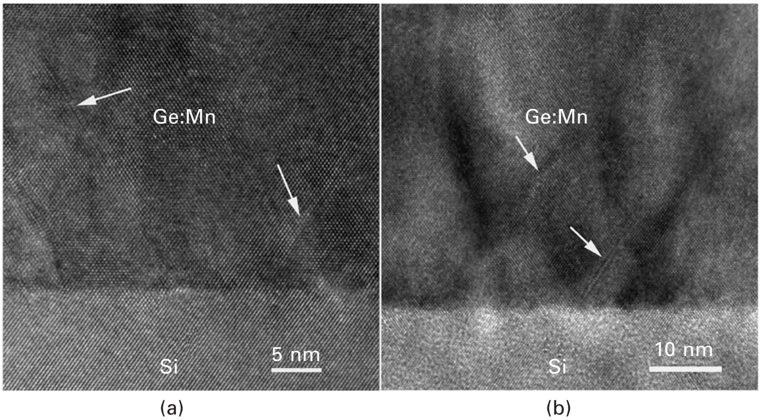
Figure 22.2 shows the TEM cross-section of as-grown and 400°C annealed samples. Stacking faults along the {111} atomic planes are clearly observed in both cases. The annealing process does not dramatically affect the stacking faults. However, the magnetic properties show a noticeable change. Figures 22.3(a) and 22.3(d) show clear hysteresis loops at 298 K for both the as-grown and annealed samples. Before annealing, the film shows a clear hysteresis at 10 K while a very small magnetic moment is observed at room temperature. After annealing, the hysteresis showed a significant change and



22.1 (a) As-grown sample at 250°C; (b) annealed sample at 400°C; (c) and (d) are corresponding Mn maps of (a) and (b), respectively; (e) HRTEM image of a typical Mn_5Ge_3 cluster; (f) HRTEM image of a typical α-Mn metallic cluster epitaxially grown on the (Mn)Ge lattice. The insets in (e) and (f) are the corresponding FFT maps of the images.

the coercive field at 10 K decreased from 430 Oe to 92 Oe. For the as-grown film, Mn_5Ge_3 formed on the surface gave a T_c of around 298 K as shown in Fig. 22.3(b). The blocking temperature was estimated to be about 250 K. After annealing, both Mn_5Ge_3 and Mn clusters led to a complex magnetic behavior, showing multiple phase transitions at different temperatures in Fig. 22.3(e). The T_c increased from 298 to over 400 K. Further investigation is needed to understand the change of the magnetic property. The Arrott's plots were made to confirm the T_c in Figs 22.3(c) and 22.3(f). The estimated Curie temperatures for the as-grown and annealed samples were obtained to be 298 K and above 400 K, respectively.

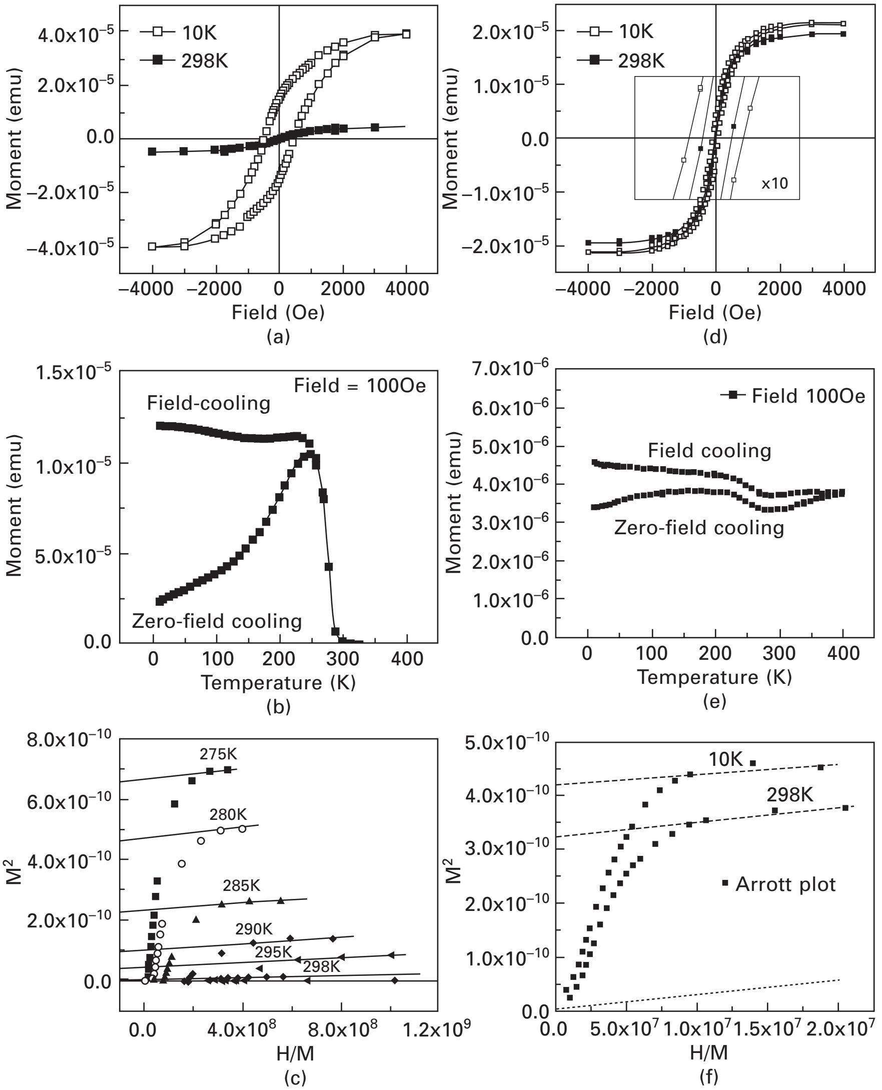
To understand the magnetic behavior of Mn in an ultra-thin Ge, a MnGe film with a thickness of about 15 nm was grown directly on Si substrate by following the same growth procedure as that for the 80 nm-thick film.



22.2 80 nm thick Mn-doped Ge grown on Si at 250°C; (a) as-grown film and (b) after 400°C annealing for 30 min. Stacking faults are marked by arrows.

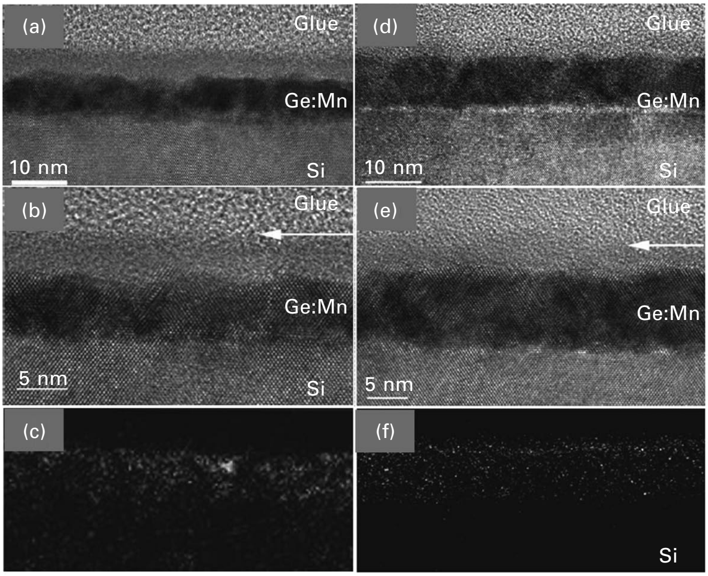
Figures 22.4(a) and 22.4(b) show typical cross-sectional TEM images and general morphologies of the as-grown and annealed samples. From these figures, a relatively uniform crystalline layer, adjacent to the Si substrates, topped with an amorphous layer, can be seen in both cases. The formation of amorphous layers is not clear yet, but it could be attributed to the growth conditions and the sample preparation for TEM when exposing to the air.

To understand the structural variation for the two cases, high-resolution TEM was carried out and typical ⟨110⟩ zone-axis HR-TEM images are shown in Figs 22.4(b) and 22.4(e) for the two cases, respectively, where the amorphous layers can be clearly seen. From these HR-TEM images, stacking faults in the crystalline layers can also be seen in both cases. No structural clusters were found through our extensive HR-TEM investigations and no secondary phases were detected by our XRD measurements (not shown here). To further understand the Mn distribution in the MnGe film (crystalline and amorphous), energy-filtered TEM was performed. Figures 22.4(c) and 22.4(f) give the EF-TEM elemental maps of Mn that correspond to Figs 22.4(a) and 22.4(d), respectively. It is of interest to note that, in the as-grown case, a higher Mn concentration is found in the amorphous layer when comparing Fig. 22.4(a) with 22.4(c). The overall concentration of Mn is low, so that the contrast in the Mn maps tends to be faint. Nevertheless, the distribution of Mn in the entire film can be clearly observed. In contrast, the Mn map (Fig. 22.4(f)) in the annealed case shows some Mn distribution in the entire film, suggesting that the annealing has not only promoted crystallization but also redistributed Mn in the entire MnGe film. Nevertheless, a higher Mn concentration can still be observed in the surface region in the annealed sample (refer to Fig. 22.4(f)). The fact that, for the as-grown sample, most



22.3 SQUID measurement of 4% Mn-doped Ge grown on Si. The left column is for the as-grown film and the right column is for the film after 400°C annealing; (a) and (d) are the hysteresis loops at 10 and 298 K; (b) and (e) are moment versus temperature curves; (c) and (f) are Arrott's plots. The dotted line in (f) shows the expected Arrott's plot at the Curie temperature.

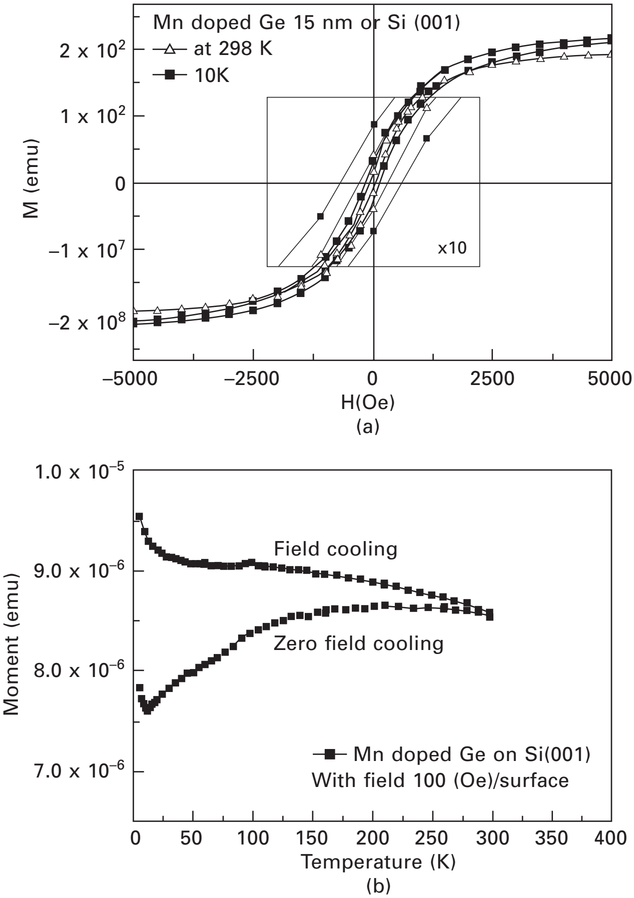
of the Mn was found in the topmost layer suggests that, during the growth of the MnGe film, a dynamic Mn diffusion has taken place, even at a relatively low growth temperature of 250°C as in our case; Mn tends to diffuse towards the top surface.



22.4 Typical TEM images of (a) the as-grown and (d) the annealed MnGe thin films with a thickness of 15 nm. The left column is for the as-grown film and the right column is for the film after 400°C annealing; (b) and (e) are the corresponding HRTEM images of (a) and (d), respectively; (c) and (f) show EFTEM elemental maps of Mn corresponding to the regions given by (a) and (d), respectively. The white arrows in (b) and (e) show the separation between the glue and the amorphous MnGe layer.

Figure 22.5(a) shows the hysteresis loops at 298 K and 10 K. A clear hysteresis at 298 K indicates room-temperature ferromagnetism. Figure 22.5(b) shows the plot of the magnetic moment as a function of temperature at a magnetic field of 100 Oe. The moment in the FC curve decreases very marginally as the temperature increases from 10 to 298 K, showing that the T_c is above the room temperature. The separation of the ZFC data from that of the FC shows a spin-glass phase. Before and after annealing, there is no clear difference observed in magnetic properties. The T_c can only be inferred to be above 400 K as shown in Fig. 22.5(b).

The difference of structural and magnetic properties between the 80 nm-thick and the 15 nm-thick MnGe thin films suggests that when it comes to the nanostructures, the strain and the Mn distribution could be different from that of the bulk films. There might be a threshold thickness for a certain Mn concentration, beyond which the nucleation of metallic clusters (such as Mn_5Ge_3 and $Mn_{11}Ge_8$) takes place and accounts for the RT ferromagnetism (T_c = 296–300 K).[83,106] It is therefore of great interest to further investigate MnGe nanostructures to explore their magnetic properties. MnGe nanodots



22.5 SQUID measurement of 4% Mn-doped Ge thin film grown on Si with a thickness of 15 nm: (a) hysteresis loops measured at 10 and 298 K; (b) magnetic moment versus temperature measured with a magnetic field of 100 Oe.

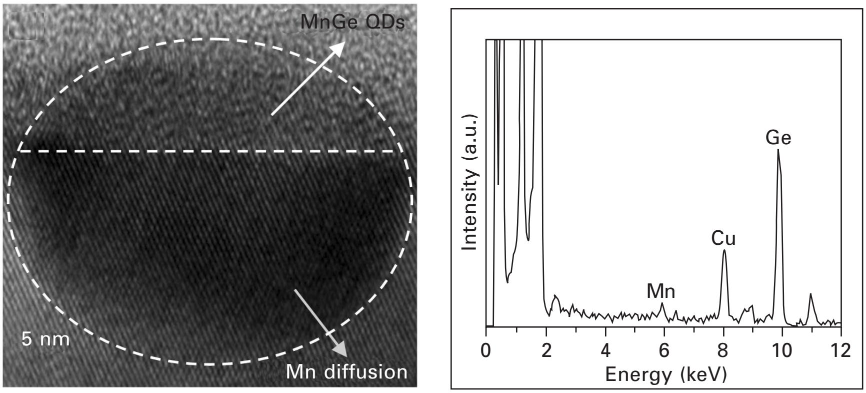
may be formed by an ion-implantation process of Mn into Ge and by a self-assembling process with MBE. Since the Mn ion implantation induces metallic clusters and defects,[91] we will only focus on the MnGe DMS QDs fabricated via MBE.

Self-assembled DMS MnGe quantum dots (QDs)

$Mn_{0.05}Ge_{0.95}$ QDs were grown on *p*-type Si substrates. Cross-section TEM was carried out to determine the structural characteristics and the Mn composition. The HR-TEM image reveals that a QD has a dome shape with a base diameter of about 30 nm and a height of about 8 nm (Fig. 22.6(a)).

The interface between the dot and the Si substrate has excellent lattice coherence. A careful inspection reveals that the dot is single-crystalline without evidence of pronounced dislocations or stacking faults. However, because of the heavy Mn doping it is possible that some amount of point defects (such as Mn interstitials) may be present inside the dot, which is beyond the detection capability of conventional TEM.

Directly underneath the $Mn_{0.05}Ge_{0.95}$ QD, Mn diffuses into Si substrate and forms a strained MnSi area, which has the same diameter as the top $Mn_{0.05}Ge_{0.95}$ QD but a height of about 16 nm (Fig. 22.6(a)). This diffusion behavior is not unusual as it was also observed in (In, Mn)As and (In, Cr)As QDs systems.[118,119] However, the migrations of Mn into the substrate make it difficult to accurately determine Mn concentration inside the dot. To address this challenge, we have performed energy dispersive X-ray spectroscopy (EDS) experiments (in a scanning TEM mode) to analyze the Mn composition at nanoscale (Fig. 22.6(b)). When the EDS was performed, electron probes several nanometers in diameter were illustrated on the $Mn_{0.05}Ge_{0.95}$ QDs and their underlying Si substrate. Consequently Si peaks are constantly observed in the EDS spectra. To estimate the Mn concentrations, we performed a quantitative analysis of atomic percentages of Mn and Ge and artificially discounted the Si peak. The EDS analysis over many QDs reveals a Mn/Ge atomic ratio of about 0.144:1. Since approximately one-third volume fraction of Mn is distributed in the MnGe QD,[40] the average Mn concentration can be estimated to be 4.8 ± 0.5%. Note that the deviation was determined by a thorough study of many MnGe QDs. Both the HR-TEM investigations



(a) (b)

22.6 Structural properties of $Mn_{0.05}Ge_{0.95}$ quantum dots grown on a p-type Si substrate. (a) High-resolution TEM cross-section image of typical $Mn_{0.05}Ge_{0.95}$ quantum dot showing the detailed lattice structure. (b) Typical EDS spectrum showing that both Mn and Ge are present in the $Mn_{0.05}Ge_{0.95}$ quantum dot. The average Mn concentration in the quantum dots was estimated to be 4.8 ± 0.5%.

and the composition analysis suggest that each individual QD is a single-crystalline DMS system.

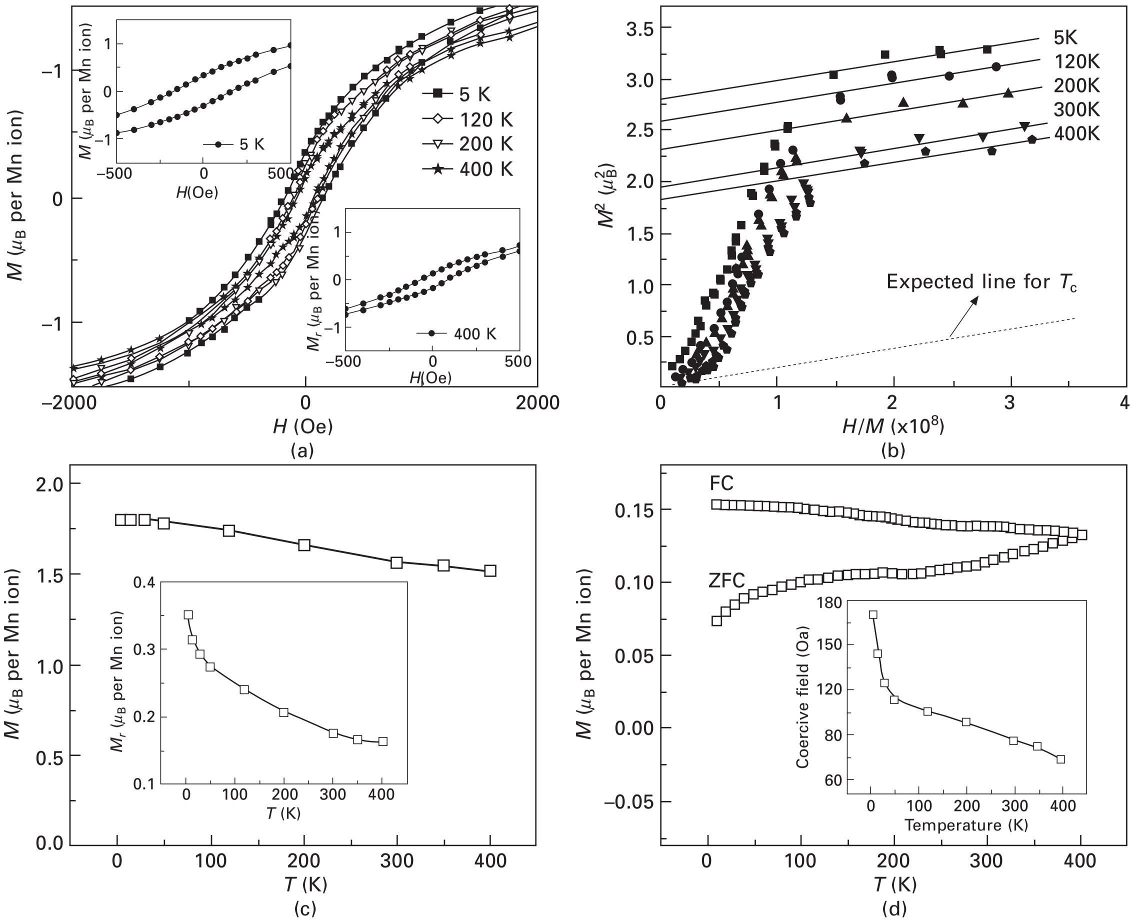
Magnetic properties were studied using a superconducting quantum interference device (SQUID) magnetometer at various temperatures. Figure 22.7(a) and the two corresponding insets show the temperature-dependent hysteresis loops when the external magnetic field is parallel to the sample surface (in-plane). The field-dependent magnetization indicates a strong ferromagnetism above 400 K. The saturation magnetic moment per Mn atom is roughly estimated to be $1.8\mu_B$ at 5 K. A fraction of roughly 60% of Mn is estimated to be activated assuming that each Mn has a moment of $3\mu_B$.[27,59] The Arrott plots[2] were also made in order to evaluate the T_c (Fig. 22.7(b)). By neglecting the high-order terms, the magnetic field can be expressed in the following equation:[120]

$$H = \frac{1}{\chi}M + \beta M^3 \quad \text{or} \quad \frac{H}{M} = \frac{1}{\chi} + \beta M^2 \qquad 22.1$$

where H is the external magnetic field, M is the magnetic moment from the sample, χ is the susceptibility, and β is a material-dependent constant.

According to Equation 22.1, M^2 can be plotted as a function of H/M. When H/M is extrapolated to $M^2 = 0$, the intercept on the H/M axis gives $1/\chi$. The T_c for this material can be obtained when $1/\chi$ vanishes. Figure 22.7(b) shows the Arrott plots at several temperatures. It can be seen that even at 400 K the intercept ($1/\chi$) on the H/M axis does not vanish, which means that the susceptibility still has a finite value and the T_c has not been reached yet. By using the slope obtained at 400 K, a dashed line can be drawn as shown in Fig. 22.7(b), in which a T_c is projected to be beyond 400 K. This is in good agreement with the data from the hysteresis loops showing the magnetic order above 400 K. Figure 22.7(c) and the corresponding inset show the temperature-dependent saturation and remnant moments per Mn ion, respectively. Both demonstrate weak temperature dependences, and substantial amounts of magnetization moments remain even at 400 K.

ZFC and FC magnetizations were measured with a magnetic field of 100 Oe as shown in Fig. 22.7(d). The magnetic moments do not drop to zero, suggesting a high T_c beyond 400 K, which is in good agreement with the Arrott plots in Fig. 22.7(b). From these two curves, one can also infer the formation of a single phase in this material system, i.e., DMS QDs. The wide separation of the ZFC and FC curves in the temperature range of 5 to 400 K shows the irreversibility of susceptibilities, possibly arising from strain-induced anisotropy as a large lattice mismatch exists between Si and Ge.[62] The temperature-dependent coercivity is shown in Fig. 22.7(d) inset. The coercivity decreases from 170 Oe (at 5 K) to 73 Oe (at 400 K). The small coercivity in the entire temperature range measured features a soft ferromagnetism which originates from Mn ions diluted in the Ge matrix.[44]



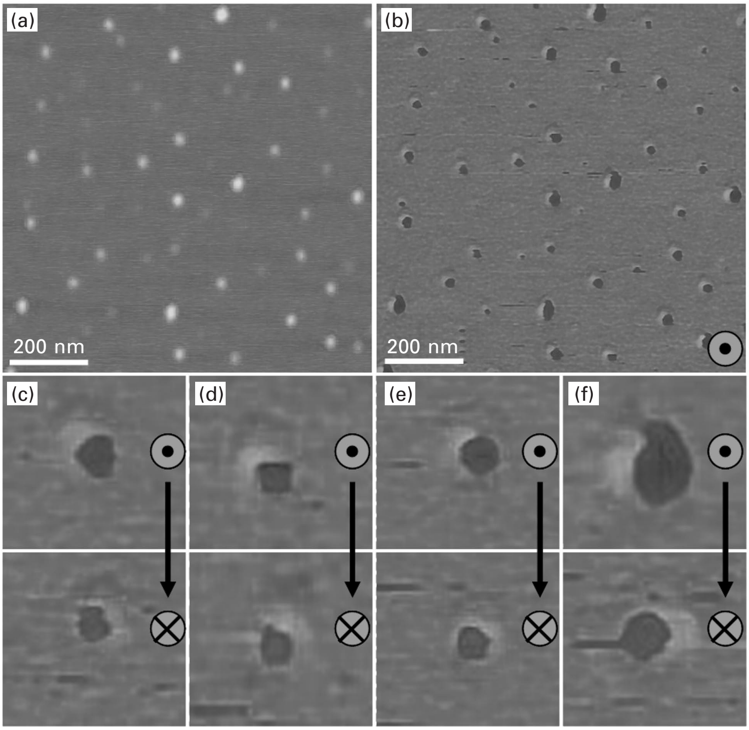
22.7 Magnetic properties of the $Mn_{0.05}Ge_{0.95}$ quantum dots grown on a p-type Si substrate. (a) Hysteresis loops measured at different temperatures from 5 to 400 K. (b) The Arrott plots were made to obtain the Curie temperature. Consistent with (a), the Curie temperature is projected to be above 400 K. (c) Temperature dependence of saturation moments; the inset gives the remnant moments with respect to temperature. (d) Zero-field cooled and field cooled magnetizations of quantum dots with a magnetic field of 100 Oe; the inset shows the coercivity values at different temperatures. The external magnetic field is in parallel with the sample surface.

The above magnetic properties support the fact that the $Mn_{0.05}Ge_{0.95}$ QDs exhibit a DMS-type ferromagnetic order.

It is well known that the direct interaction (superexchange) between *d*-shells of the adjacent Mn atoms leads to an antiferromagnetic configuration of the *d*-shell spins.[8] Therefore, the presence of Mn–Mn clusters would lead to the antiferromagnetism; but when this pair is surrounded by other Mn atoms, the ferri- or ferromagnetic coupling may appear. However, the extensive high-resolution TEM did not reveal observable Mn clusters, although such a possibility cannot be completely excluded (Fig. 22.6). Thus, the origin of strong ferromagnetism possibly comes from the $Mn_{0.05}Ge_{0.95}$ QDs via the hole mediation process. In general, when Mn dopants are introduced into Ge, they produce two acceptor levels, i.e., 0.16 eV and 0.37 eV above the valence band.[121] When the doping concentration becomes high, the impurity bands could be developed (instead of being separate acceptor levels). However, theoretically the formation of such impurity bands may not affect the physics of ferromagnetism.[122]

In the following, we will analyze the Mn doping behavior by considering a single acceptor level of 0.16 eV in an approximate form. At a sufficiently high temperature, Mn acceptors can be activated to generate holes. These holes are itinerant among a number of Mn ions and align Mn local spins along one direction via the *p-d* exchange coupling to reduce the total energy of the system, resulting in the hole-mediated ferromagnetism.[123] Experimentally this phenomenon has been widely observed in several material systems, such as group III–V (GaAs and InAs),[2,124] group II–VI (ZnO),[125] and group IV (Ge).[27,39] While most research focuses on bulk DMS, of particular interest is to study the origin of ferromagnetism in quantum structures such as (In, Mn) As and MnGe QDs. Previously, Chen *et al.*[39] and others[118] have suggested that the carrier confinement in a DMS QD can strengthen hole localization and subsequently enhance the thermal stability of magnetic polarons, giving rise to a higher T_c than for bulk films.

Atomic force microscopy (AFM) and magnetic force microscopy (MFM) measurements were carried out to investigate the morphology and ferromagnetism of $Mn_{0.05}Ge_{0.95}$ QDs at 320 K, respectively. The average dot size is 50 nm in base diameter and 6 nm in height. The dot density is about 6×10^9 cm^{-2} (Fig. 22.8(a)). The corresponding MFM image was taken by lifting up the MFM probe 25 nm above the topographic height of the sample in a phase detection mode (Fig. 22.8(b)). The appearance of bright and dark areas in the MFM image clearly shows the formation of magnetic domains in the $Mn_{0.05}Ge_{0.95}$ QDs, which is similar to that in (In, Mn)As DMS QDs.[126] Figures 22.8(c)–22.8(f) show enlarged MFM images of several individual $Mn_{0.05}Ge_{0.95}$ QDs. By reversing the tip magnetization, opposite contrast was observed for each dot, indicating that the magnetic signals originated from the top $Mn_{0.05}Ge_{0.95}$ QDs. Note that each QD is a single domain ‘particle’.



22.8 AFM and MFM images of the $Mn_{0.05}Ge_{0.95}$ quantum dots measured at 320 K. (a) Typical AFM image of $Mn_{0.05}Ge_{0.95}$ quantum dots; (b) corresponding MFM image with the tip magnetization pointing toward the sample; (c)–(f), enlarged MFM images of individual quantum dots taken from (b). From these MFM measurements, opposite contrasts were observed when applying an opposite magnetization to the tip. The results suggest that the magnetic signals come from the $Mn_{0.05}Ge_{0.95}$ quantum dots. The formation of Mn_5Ge_3 and $Mn_{11}Ge_8$ can be excluded because of their low Curie temperatures of 296–300 K.

During the magnetization process, the domain would rotate preferentially along the magnetic field to produce net magnetization moments. Since the experiments were performed at 320 K, the formation of metallic phases such as Mn_5Ge_3 and $Mn_{11}Ge_8$ can be easily ruled out because they have low Curie temperatures of 296–300 K. Overall, the above MFM results agree well with the TEM observations and the ferromagnetic order at high temperature obtained in the SQUID measurements.

22.3.3 Discussion

We have reviewed the growth of MnGe by MBE, ion implantation, and bulk sintering, emphasizing the growth temperature effect on the phase formation and magnetic properties. While the current motivation is driven by the prospect of enhancing Curie temperatures, the experiments show a strong tendency to metallic precipitates developed in bulk and relatively thick MnGe crystals. The fundamental understanding here points to a spin-glass-like feature in MnGe due to the formation of various lattice coherent nanostructures and metallic precipitates, which exhibit magnetic blocking behaviors under different temperature regimes. The challenge to eliminate these clusters seems to be overwhelming and nearly impossible. However, when the material dimension decreases, it evidences a change of material structure, being free of precipitates, and a significant increase of Curie temperature over 400 K. The preliminary explanation for these behaviors can be attributed to the carrier confinement in a DMS QD which strengthens hole localization and subsequently enhances the thermal stability of magnetic polarons, giving rise to a higher T_c than those of bulk films. Nevertheless, the experimental data suggest that with nanoscale structures, the quantum confinement effect comes into being, which significantly influences the exchange coupling between the confined holes and the localized Mn^{2+}. Future progress in this direction will involve a detailed theoretical treatise to understand the quantum confinement effect on magnetic properties in a quantitative picture.

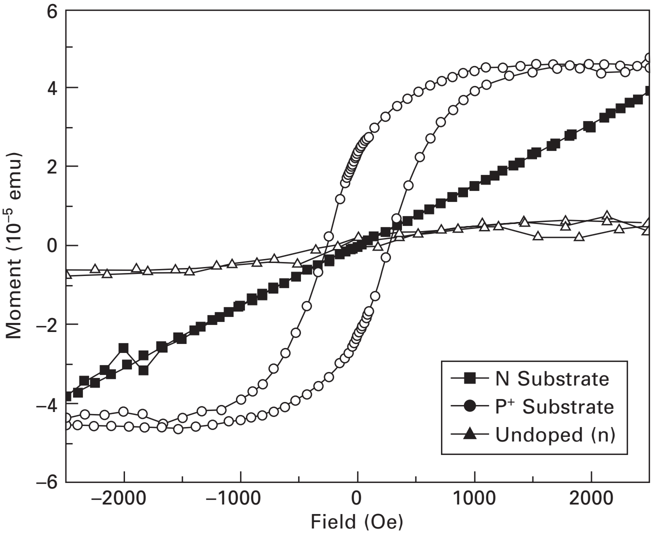
22.4 Electric field-controlled ferromagnetism

22.4.1 Introduction

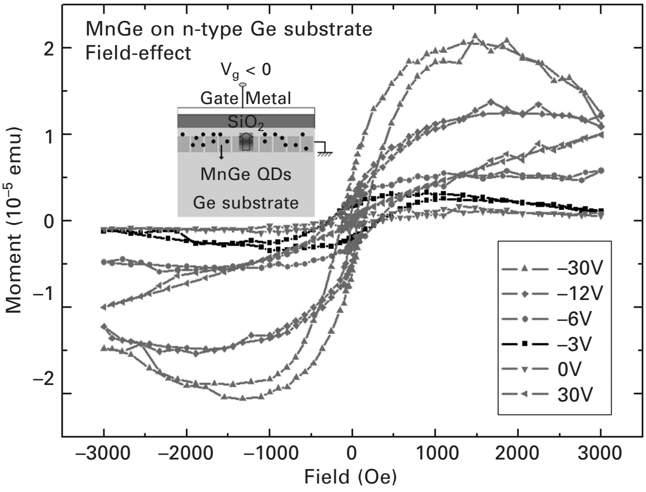
Electric field control of ferromagnetism has the potential to realize spin field-effect transistors (spin FETs) and nonvolatile spin logic devices via the carrier-mediated effect.[127,128] With the manipulation of carrier spins, a new generation of nonvolatile (green) computing systems could be eventually developed for many low-power-dissipation applications in all fields, including sensor networks, health monitoring, information systems, sustainable wireless systems, etc. Since Datta and Das[6] first introduced the concept of spin FETs in 1990, enormous efforts have been dedicated to creating a device wherein the carrier transport is modulated by electrostatic control of carrier spins.[2,20,22,23,129–133] To understand and exploit this controllability, several theoretical models were proposed to explain the ferromagnetic coupling in DMS on a microscopic scale (also elaborated in Section 22.2): Zener–kinetic exchange,[10] double-exchange,[134] and RKKY interaction.[135,136] These models share a common feature that a spontaneous ferromagnetic order is carrier-mediated through the increase of carrier concentration. One of the major challenges, however, is to seek an ideal material with room-temperature controllable spin states.[21,22,137] In recent years, DMSs have emerged to become one of the promising candidates since they could possibly offer a high T_c in excess of 300 K.[134] The demonstration of carrier-medicated ferromagnetism involving correlated electron/hole systems leads to a para- to ferro-magnetism phase transition.[27,39,128,134] In principle, the collective alignment of spin states in these DMSs can be manipulated by the modulation of carrier concentrations through gate biasing in a FET structure.[2,137] For this kind of spin FETs, the 'source' and 'drain' may be completed through 'nanomagnets', which are in turn controlled by the gate; and no carrier transport is needed. Clearly, one may also involve the control of source–drain conductance by gate-voltage-induced precession of injected spins (from the source). Since the early 2000s, significant progress has been achieved,[2,23,128,132] in which the ferromagnetism of an (In, Mn)As channel layer could be effectively turned on and off via electric fields in a gated FET. Such extraordinary field-modulated ferromagnetism immediately rendered possible the development of future spintronic devices. However, the manipulation of ferromagnetism was limited because of the low T_c of the Mn-doped III–V materials.[138] Therefore, a search for new DMS materials with $T_c > 300$ K and carrier-mediated ferromagnetism becomes a current global challenge.[10,137]

22.4.2 Electric field-controlled ferromagnetism in Mn-implanted Ge dots

In parallel with the experimental efforts on field-controlled ferromagnetism in III–V material systems, we have previously investigated the magnetic properties of Mn_xGe_{1-x} nanostructures by ion implantation of Mn.[39] We demonstrated the hole-mediated effect in Mn-implanted nanodot structures and the modulation of their ferromagnetism by applying gate biases in MOS capacitors at a low temperature of 10 K. Figure 22.9 shows the magnetic properties of samples with Mn implantation into Ge via the nano-patterned SiO_2 mask on different doped substrates. It is observed that the Mn-implanted p^+ Ge (10^{19} cm^{-3}) has stronger ferromagnetic signals than that of n-type Ge (10^{17} cm^{-3}), which may indicate a hole-mediated effect. Note that we cannot completely exclude the possibility of the formation of dopant complexes, such as B–Mn and Sb–Mn in the p^+ and n-type Ge, respectively. Taking advantage of this effect, we further fabricated a MOS capacitor device with Mn-implanted Ge as the channel layer. Figure 22.10. shows the magnetic moment versus the applied field for the MOS structure on the n-type substrate under various gate biases at 10 K. When the voltage is applied from −6 to −30 V, the hole concentration in the MnGe channel increases and the magnetic hysteresis becomes larger. At 0 V bias, only a line with a small slope and a small negligible loop are seen; this is because a low density of excess holes is present. When a positive voltage is applied to the gate, holes are partially depleted, resulting in a decrease of the channel magnetization



22.9 The magnetic properties of Mn implanted Ge via a nano-patterned SiO_2 mask. The hole-mediated effect can be clearly observed in the Mn implanted p-type Ge sample.



22.10 Magnetic moment vs applied field under various biases from −30 to 30 V at 10 K. The obtained hysteresis is the largest under −30 V gate voltage, and the ferromagnetic phase is turned off under +30 V gate bias. The inset shows the spin-gated structure used to investigate *C–V* and gate bias-dependent magnetic properties.

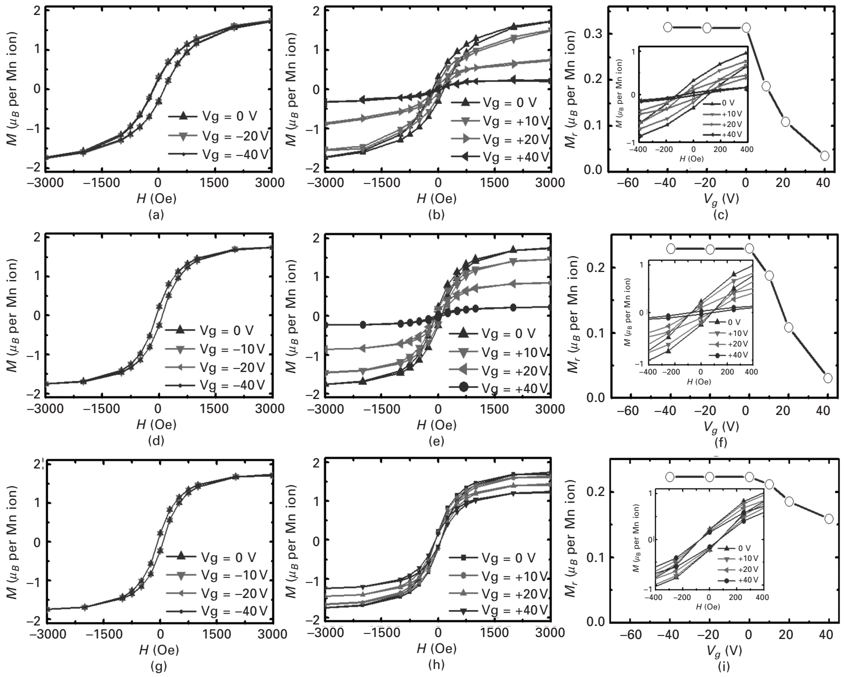
and in the disappearance of magnetic hysteresis. These experimental results demonstrated the hole-mediated effect in MnGe, consistent with the finding by Park *et al.*[27] However, metallic precipitates such as Mn_5Ge_3 and $Mn_{11}Ge_8$

and implantation damage were found in these nanostructures, which made the system rather complex and might have partly jeopardized the hole-mediated effect.

22.4.3 Electric field-controlled ferromagnetism in self-assembled DMS MnGe QDs

Since the self-assembled MnGe QDs have a high T_c above 400 K and are free of metallic phases (Mn_5Ge_3 and $Mn_{11}Ge_8$), they can be potentially used as a channel layer in a MOS device to study the modulation of ferromagnetism by an electric field. The device structure consists of a metal gate (Au, 200 nm), Al_2O_3 (40 nm), $Mn_{0.05}Ge_{0.95}$ QDs (6 nm), a wetting layer (<0.6 nm), a *p*-type Si substrate (1×10^{18} cm^{-3}), and a back metal contact (Au, 200 nm). A relatively thick dielectric of Al_2O_3 (40 nm) is employed to ensure a small leakage current to a level below 10^{-6}A/cm^2.

Figure 22.11 shows the electric field-controlled ferromagnetism performed at 50, 77, and 100 K, corresponding to (a)–(c), (d)–(f), and (g)–(i), respectively. Due to the similarity of the data, we take 77 K as an example to describe the device operation in Figs 22.11(d)–22.11(f). Figures 22.11(d) and 22.11(e) show the hysteresis loops by SQUID with negative and positive biases on the MOS gate at 77 K, respectively. Under a negative bias, the holes are attracted into the channel of the device (accumulation). In this circumstance, however, the hysteresis loop does not show a remarkable change (Fig. 22.11(d)). This could be explained by the fact that even at 0 V, the QD device has already accumulated enough holes to induce ferromagnetism. In other words, the hole-mediated effect is sufficient to align a majority of the activated Mn ions along one direction in each individual QD. Further increasing negative bias does not change the hole concentrations much. On the contrary, with the positive bias, a large quantity of holes is depleted into the *p*-type Si so that the hole-mediated effect is notably reduced. In fact, the surface of the device can have a high concentration of electrons if the leakage is limited (inversion). As a result, the Mn ions start to misalign because of the lack of holes. The saturation moment per Mn ion decreases more than 10 times as the gate bias increases from 0 to +40 V (Fig. 22.11(e)). It should be noted that, at +40 V, the saturation and remnant moments of the $Mn_{0.05}Ge_{0.95}$ QDs become fairly weak, resembling a 'paramagnetic-like' state. Figure 22.11(f) summarizes the change of remnant moments as a function of the gate voltage. A similar trend in the carrier density as a function of the gate bias was observed from CV curves (not shown here), suggesting a strong correlation between the hole concentrations and the ferromagnetism, i.e. the hole-mediated effect. The inset in Fig. 22.11(f) displays an enlarged picture to clearly show the change of remnant moments with respect to the gate bias. By increasing the measurement temperature to 100 K (Figs 22.11(g)–22.11(i)), the modulation



22.11 Control of ferromagnetism of $Mn_{0.05}Ge_{0.95}$ quantum dots by applying an electric field at 50 K ((a)–(c)), 77 K ((d)–(f)) and 100 K ((g)–(i)). (a), (d) and (g): Hysteresis loops with zero and negative bias of −10, −20 and −40 V on the gate; (b), (e) and (h): hysteresis loops with zero and positive bias of +10, +20 and +40 V; (c), (f) and (i): remnant moments with respect to the gate bias. Insets of (c), (f) and (i) are enlarged figures from the central part of (b), (e) and (h) to clearly show the change of remnant moments, respectively. It is found that both the saturation and remnant moments can be manipulated by applying biases on the MOS gate at 50, 77 and 100 K.

of the ferromagnetism becomes less pronounced compared to those at 50 and 77 K due to the increased leakage current in our MOS devices. The above results evidently demonstrate that the hole-mediated effect does exist in this material system.

22.4.4 Discussion

Electrical field-controlled ferromagnetism has been successfully demonstrated in Mn-doped III–V DMSs.[2,23,128,132,137] It is well established that the variation of hole concentrations renders such controllability.[2] The low T_c of III–V DMS, however, presents a challenge to further realizing room-temperature-controlled ferromagnetism.[138] Alternatively, the recent experiments on MnGe QDs show a high T_c above room temperature and gate-modulated ferromagnetism over 100 K.[40] While the leakage current suppresses the gate modulation, room-temperature-controlled ferromagnetism would not be impossible because of the high T_c of this system. Research in this direction may open up a pathway for achieving room-temperature Ge-based (and other) spin-FETs and spin logic devices. These spintronic devices could potentially replace conventional FETs with lower power consumption and provide additional new device functionalities, which is beyond today's mainstream CMOS technology of microelectronics.

22.5 Conclusion and future trends

This chapter is a review of theoretical and experimental progress that has been achieved in understanding ferromagnetism and related electronic properties in MnGe DMSs. Interest in DMS ferromagnetism is motivated by the possibility of engineering systems that combine many of the technologically useful features of ferromagnetic and semiconducting materials. This goal has been achieved to an impressive degree in (III, Mn)V DMSs, and further progress can be anticipated in the future. However, due to the low T_c of (III, Mn) V DMSs, spintronics research seems to have reached a critical bottleneck, in which achieving a high T_c DMS becomes an intriguing and challenging task. Fortunately, the MnGe material system offers a possible route towards higher T_c. The structural and magnetic properties can be adjusted simply by modifying the system dimensions from 3-D thin films to 0-D quantum dots. A high T_c in excess of 400 K can be obtained and is presumably attributed to the quantum confinement effect, which strengthens the hybridization between the localized Mn impurities and the it in erant holes. Bound magnetic polarons may also exist, since this system falls into a regime where Mn concentrations are much larger than those of holes. Nevertheless, the high T_c in the low-dimensional system suggests that the quantum structure exhibits extraordinary properties that significantly differ from those of bulk films. Future progress would rely on the precise theoretical understanding of the quantum confinement effect on ferromagnetism.

It is well known that achieving functioning spintronic devices working at ambient temperatures requires the following criteria: (1) the ferromagnetic transition temperature should safely exceed room temperature, (2) the mobile charge carriers should respond strongly to changes in the ordered magnetic state, and (3) the material should retain fundamental semiconductor characteristics, including sensitivity to doping and light, and electric fields produced by gate charges. For more than a decade, these three key issues have been the focus of intense experimental and theoretical research. Progress has also been made in achieving field-controlled ferromagnetism in (III, Mn)As systems, even though controllability remains at low temperatures because of low T_c. Therefore, the critical challenge now is either to continue increasing T_c in (III, Mn)As or to look for a new DMS system with both high Curie temperature ($T_c \gg 300$ K) and field-controlled ferromagnetism to satisfy all these three criteria. The experimental results on field-controlled ferromagnetism in MnGe QDs suggest that the ferromagnetism in this system sensitively responds to the electrical field via the hole-mediated effect, in a similar way to that in (III, Mn)As systems. Therefore, with a much higher T_c compared with III–V DMS, MnGe nanostructures could become one of the most promising candidates to achieve room-temperature operation. We also hope to encourage more research efforts in this direction because of the excellent compatibility of Ge with the current CMOS technology.

Besides DMS-based spintronics devices, another type of spin-devices based on Si and Ge becomes possible by forming DO_3-type ferromagnetic silicide, the details and magnetic properties of which are described in Chapter 9.

22.6 References

1 Prinz, G.A., Magnetoelectronics. *Science* 282 (5394), 1660–1663 (1998).

2 Ohno, H. *et al.*, Electric-field control of ferromagnetism. *Nature* 408 (6815), 944–946 (2000).

3 ITRS, Emerging Research Materials. *International Technology Roadmap for Semiconductors* (2007 edition).

4 Holub, M. *et al.*, Spin-polarized light-emitting diodes and lasers. *Journal of Physics D: Applied Physics* 11, R179 (2007).

5 Truong, V.G. *et al.*, High speed pulsed electrical spin injection in spin-light emitting diode. *Applied Physics Letters* 94 (14), 141109 (2009).

6 Datta, S. & Das, B., Electronic analog of the electro-optic modulator. *Applied Physics Letters* 56 (7), 665–667 (1990).

7 Sugahara, S. & Tanaka, M., A spin metal–oxide–semiconductor field-effect transistor using half-metallic-ferromagnet contacts for the source and drain. *Applied Physics Letters* 84 (13), 2307–2309 (2004).

8 Liu, C., Yun, F., & Morkoç, H., Ferromagnetism of ZnO and GaN: a review. *Journal of Materials Science: Materials in Electronics* 16 (9), 555–597 (2005).

9 Das Sarma, S., Hwang, E.H. & Kaminski, A., How to make semiconductors ferromagnetic: a first course on spintronics. *Solid State Communications* 127, 99–107 (2003).

10 Jungwirth, T., Sinova, J., Masek, J., Kucera, J., & MacDonald, A.H., Theory of ferromagnetic (III,Mn)V semiconductors. *Reviews of Modern Physics* 78 (3), 809 (2006).

11 Ohno, H. *et al.*, (Ga,Mn)As: a new diluted magnetic semiconductor based on GaAs. *Applied Physics Letters* 69 (3), 363–365 (1996).

12 Novak, V. *et al.*, Curie point singularity in the temperature derivative of resistivity in (Ga,Mn)As. *Physical Review Letters* 101 (7), 077201 (2008).

13 Wang, M. *et al.*, Achieving high Curie temperature in (Ga,Mn)As. *Applied Physics Letters* 93 (13), 132103 (2008).

14 Olejnik, K. *et al.*, Enhanced annealing, high Curie temperature, and low-voltage

gating in (Ga,Mn)As: a surface oxide control study. *Physical Review B (Condensed Matter and Materials Physics)* 78 (5), 054403 (2008).

15 Chen, L. *et al.*, Low-temperature magnetotransport behaviors of heavily Mn-doped (Ga,Mn)As films with high ferromagnetic transition temperature. *Applied Physics Letters* 95 (18), 182505 (2009).

16 Nazmul, A.M., Amemiya, T., Shuto, Y., Sugahara, S., & Tanaka, M., High temperature ferromagnetism in GaAs-based heterostructures with Mn delta doping. *Physical Review Letters* 95 (1), 017201 (2005).

17 Nazmul, A.M., Amemiya, T., Shuto, Y., Sugahara, S., & Tanaka, M., Erratum: High temperature ferromagnetism in GaAs-based heterostructures with Mn delta doping [*Phys. Rev. Lett.* 95, 017201 (2005)]. *Physical Review Letters* 96 (14), 149901 (2006).

18 Kulbachinskii, V.A., Gurin, P.V., Vikhrova, O.V., Danilov, Y.A. & Zvonkov, B.N., Ferromagnetism and magnetotransport in GaAs structures with InAs quantum dot layer or $In_xGa_{1-x}As$ quantum well delta-doped with Mn and C. *Journal of Physics: Conference Series* 100 (04025) (2008).

19 Wurstbauer, U. & Wegscheider, W., Magnetic ordering effects in a Mn-modulation-doped high mobility two-dimensional hole system. *Physical Review B (Condensed Matter and Materials Physics)* 79 (15), 155444 (2009).

20 Nepal, N. *et al.*, Electric field control of room temperature ferromagnetism in III–N dilute magnetic semiconductor films. *Applied Physics Letters* 94 (13), 132505 (2009).

21 Kanki, T., Tanaka, H., & Kawai, T., Electric control of room temperature ferromagnetism in a $Pb(Zr_{0.2}Ti_{0.8})O_3/La_{0.85}Ba_{0.15}MnO_3$ field-effect transistor. *Applied Physics Letters* 89 (24), 242506 (2006).

22 Philip, J. *et al.*, Carrier-controlled ferromagnetism in transparent oxide semiconductors. *Nature Materials* 5 (4), 298–304 (2006).

23 Chiba, D., Matsukura, F., & Ohno, H., Electric-field control of ferromagnetism in (Ga,Mn)As. *Applied Physics Letters* 89 (16), 162505 (2006).

24 Kane, B.E., A silicon-based nuclear spin quantum computer. *Nature* 393 (6681), 133–137 (1998).

25 Vrijen, R. *et al.*, Electron-spin-resonance transistors for quantum computing in silicon–germanium heterostructures. *Physical Review A* 62 (1), 012306 (2000).

26 Miyoshi, T., Matsui, T., Tsuda, H., Mabuchi, H., & Morii, K., Magnetic and electric properties of Mn_5Ge_3/Ge nanostructured films. *Journal of Applied Physics* 85, 5372–5374 (1999).

27 Park, Y.D. *et al.*, A group-IV ferromagnetic semiconductor: Mn_xGe1-_x. *Science* 295 (5555), 651–654 (2002).

28 Tsui, F. *et al.*, Novel germanium-based magnetic semiconductors. *Physical Review Letters* 91 (17), 177203 (2003).

29 D'Orazio, F., Lucari, F., Pinto, N., Morresi, L., & Murri, R., Toward room temperature ferromagnetism of Ge:Mn systems. *Journal of Magnetism and Magnetic Materials* 272–276 (part 3), 2006–2007 (2004).

30 Kazakova, O., Kulkarni, J.S., Holmes, J.D., & Demokritov, S.O., Room-temperature ferromagnetism in $Ge_{1-x}Mn_x$ nanowires. *Physical Review B* 72 (9), 094415 (2005).

31 Li, A.P. *et al.*, Magnetism in Mn_xGe_{1-x} semiconductors mediated by impurity band carriers. *Phys Rev B* 72 (19), 195205 (2005).

32 Demidov, E. *et al.*, Ferromagnetism in epitaxial germanium and silicon layers

supersaturated with manganese and iron impurities. *JETP Letters* 83 (12), 568–571 (2006).

33 Jamet, M. *et al.*, High-Curie-temperature ferromagnetism in self-organized $Ge_{1-x}Mn_x$ nanocolumns. *Nature Materials* 5 (8), 653–659 (2006).

34 Collins, B.A., Chu, Y.S., He, L., Zhong, Y., & Tsui, F., Dopant stability and strain states in Co and Mn-doped Ge (001) epitaxial films. *Physical Review B (Condensed Matter and Materials Physics)* 77 (19), 193301 (2008).

35 Ogawa, M. *et al.*, Mn distribution behaviors and magnetic properties of GeMn films grown on Si (001) substrates. *Journal of Crystal Growth* 311 (7), 2147–2150 (2009).

36 Tsuchida, R., Asubar, J.T., Jinbo, Y., & Uchitomi, N., MBE growth and properties of GeMn thin films on (001 GaAs. *Journal of Crystal Growth* 311 (3), 937–940 (2009).

37 Cho, S. *et al.*, Ferromagnetism in Mn-doped Ge. *Physical Review B* 66 (3), 033303 (2002).

38 Bolduc, M. *et al.*, Above room temperature ferromagnetism in Mn-ion implanted Si. *Physical Review B* 71 (3), 033302 (2005).

39 Chen, J., Wang, K.L., & Galatsis, K., Electrical field control magnetic phase transition in nanostructured Mn_xGe_{1-x}. *Applied Physics Letters* 90 (1), 012501 (2007).

40 Xiu, F. *et al.*, Electric-field-controlled ferromagnetism in high-Curie-temperature $Mn_{0.05}Ge_{0.95}$ quantum dots. *Nature Materials* 9, 337–344 (2010).

41 Zhu, W., Zhang, Z., & Kaxiras, E., Dopant-assisted concentration enhancement of substitutional Mn in Si and Ge. *Physical Review Letters* 100 (2), 027205 (2008).

42 Li, A.P. *et al.*, Dopant segregation and giant magnetoresistance in manganese-doped germanium. *Physical Review B (Condensed Matter and Materials Physics)* 75 (20), 201201 (2007).

43 van der Meulen, M.I. *et al.*, Single crystalline $Ge_{1-x}Mn_x$ nanowires as building blocks for nanoelectronics. *Nano Letters* 9 (1), 50–56 (2008).

44 Kazakova, O., Kulkarni, J.S., Holmes, J.D., & Demokritov, S.O., Room-temperature ferromagnetism in $Ge_{1-x}Mn_x$ nanowires. *Physical Review B* 72 (9), 094415 (2005).

45 Majumdar, S., Mandal, S., Das, A.K., & Ray, S.K., Synthesis and temperature dependent photoluminescence properties of Mn doped Ge nanowires. *Journal of Applied Physics* 105 (2), 024302 (2009).

46 Seong, H.-K. *et al.*, Magnetic and electrical properties of single-crystalline Mn-doped Ge nanowires. *Journal of Physical Chemistry C* 113 (25), 10847–10852 (2009).

47 Yoon, I.T. *et al.*, Magnetic and transport properties of Mn-implanted Ge/Si quantum dots. *Solid State Communications* 140 (3–4), 185–187 (2006).

48 Wang, H.-Y. & Qian, M.C., 2006 (unpublished).

49 Chang, Y.Q. *et al.*, Synthesis, optical, and magnetic properties of diluted magnetic semiconductor $Zn_{1-x}Mn_xO$ nanowires via vapor phase growth. *Applied Physics Letters* 83 (19), 4020–4022 (2003).

50 Cui, J.B. & Gibson, U.J., Electrodeposition and room temperature ferromagnetic anisotropy of Co and Ni-doped ZnO nanowire arrays. *Applied Physics Letters* 87 (13), 133108 (2005).

51 Baik, J.M. & Lee, J.L., Fabrication of vertically well-aligned (Zn,Mn)O nanorods

with room temperature ferromagnetism. *Advances in Materials* 17 (22), 2745–2748 (2005).

52 Brieler, F.J. *et al.*, Formation of $Zn_{1-x}Mn_xS$ nanowires within mesoporous silica of different pore sizes. *Journal of the American Chemical Society* 126 (3), 797–807 (2004).

53 Radovanovic, P.V., Barrelet, C.J., Gradecak, S., Qian, F., & Lieber, C.M., General synthesis of manganese-doped II–VI and III–V semiconductor nanowires. *Nano Letters* 5 (7), 1407–1411 (2005).

54 Choi, H.J. *et al.*, Single-crystalline diluted magnetic semiconductor GaN:Mn nanowires. *Advances in Materials* 17 (11), 1351-+ (2005).

55 Jeon, H.C., Chung, K.J., Chung, K.J., Kang, T.W., & Kim, T.W., Enhancement of the ferromagnetic transition temperature in self-assembled $(Ga_{1-x}Mn_x)As$ quantum wires. *Japanese Journal of Applied Physics Part 2* 43 (7B), L963–L965 (2004).

56 Zener, C., Interaction between the d shells in the transition metals. *Physical Review* 81 (3), 440 (1951).

57 Zhao, Y.-J., Shishidou, T., & Freeman, A.J., Ruderman–Kittel–Kasuya–Yosida-like ferromagnetism in Mn_xGe_{1-x}. *Physical Review Letters* 90 (4), 047204 (2003).

58 Das Sarma, S., Hwang, E.H., & Kaminski, A., Temperature-dependent magnetization in diluted magnetic semiconductors. *Physical Review B* 67 (15), 155201 (2003).

59 Stroppa, A., Picozzi, S., Continenza, A., & Freeman, A.J., Electronic structure and ferromagnetism of Mn-doped group-IV semiconductors. *Physical Review B* 68 (15), 155203 (2003).

60 Zener, C., Interaction between the d-shells in the transition metals. II. Ferromagnetic compounds of manganese with perovskite structure. *Physical Review* 82 (3), 403 (1951).

61 Akai, H., Ferromagnetism and its stability in the diluted magnetic semiconductor (In, Mn)As. *Physical Review Letters* 81 (14), 3002 (1998).

62 Dietl, T., Ohno, H., & Matsukura, F., Hole-mediated ferromagnetism in tetrahedrally coordinated semiconductors. *Physical Review B* 63 (19), 195205 (2001).

63 Picozzi, S. & Lezaic, M., Ab-initio study of exchange constants and electronic structure in diluted magnetic group-IV semiconductors. *New Journal of Physics* (5), 055017 (2008).

64 Luo, X., Zhang, S.B., & Wei, S.-H., Theory of Mn supersaturation in Si and Ge. *Physical Review B* 70 (3), 033308 (2004).

65 Erwin, S.C. & Petukhov, A.G., Self-compensation in manganese-doped ferromagnetic semiconductors. *Physical Review Letters* 89 (22), 227201 (2002).

66 Zhu, W., Weitering, H.H., Wang, E.G., Kaxiras, E., & Zhang, Z., Contrasting growth modes of Mn on Ge(100) and Ge(111) surfaces: Subsurface segregation versus intermixing. *Physical Review Letters* 93 (12), 126102 (2004).

67 Continenza, A., Profeta, G., & Picozzi, S., Transition metal impurities in Ge: Chemical trends and codoping studied by electronic structure calculations. *Physical Review B (Condensed Matter and Materials Physics)* 73 (3), 035212 (2006).

68 Chen, H., Zhu, W., Kaxiras, E., & Zhang, Z., Optimization of Mn doping in group-IV-based dilute magnetic semiconductors by electronic codopants. *Physical Review B (Condensed Matter and Materials Physics)* 79 (23), 235202 (2009).

69 Kaminski, A. & Das Sarma, S., Polaron percolation in diluted magnetic semiconductors. *Physical Review Letters* 88 (24), 247202 (2002).

70 Kaminski, A. & Das Sarma, S., Magnetic and transport percolation in diluted magnetic semiconductors. *Physical Review B* 68 (23), 235210 (2003).

71 Jaeger, C. *et al.*, Spin-glass-like behavior of Ge:Mn. *Physical Review B (Condensed Matter and Materials Physics)* 74 (4), 045330 (2006).

72 Peressi, M., Debernardi, A., Picozzi, S., Antoniella, F., & Continenza, A., Half-metallic Mn-doped Si_xGe_{1-x} alloys: a first principles study. *Computational Materials Science* 33 (1–3), 125–131 (2005).

73 Zeng, C., Zhang, Z., Benthem, K.V., Chisholm, M.F., & Weitering, H.H., Optimal doping control of magnetic semiconductors via subsurfactant epitaxy. *Physical Review Letters* 100 (6), 066101 (2008).

74 Sapra, S., Sarma, D.D., Sanvito, S., & Hill, N.A., Influence of quantum confinement on the electronic and magnetic properties of (Ga,Mn)As diluted magnetic semiconductor. *Nano Letters* 2 (6), 605–608 (2002).

75 Abolfath, R.M., Hawrylak, P & Zutik, I., Electronic states of magnetic quantum dots. *New Journal of Physics* 9 (353), 1367 (2007).

76 Govorov, A.O., Voltage-tunable ferromagnetism in semimagnetic quantum dots with few particles: Magnetic polarons and electrical capacitance. *Physical Review B* 72 (7), 075359 (2005).

77 Ahlers, S. *et al.*, Ferromagnetic Ge(Mn) nanostructures. *Physica E: Low-dimensional Systems and Nanostructures* 32 (1–2), 422–425 (2006).

78 Ahlers, S. *et al.*, Magnetic and structural properties of Ge_xMn_{1-x} films: Precipitation of intermetallic nanomagnets. *Physical Review B (Condensed Matter and Materials Physics)* 74 (21), 214411 (2006).

79 Bihler, C. *et al.*, Structural and magnetic properties of Mn_5Ge_3 clusters in a dilute magnetic germanium matrix. *Applied Physics Letters* 88 (11), 112506 (2006).

80 Bougeard, D., Ahlers, S., Trampert, A., Sircar, N., & Abstreiter, G., Clustering in a precipitate-free GeMn magnetic semiconductor. *Physical Review Letters* 97 (23), 237202 (2006).

81 Devillers, T. *et al.*, Structure and magnetism of self-organized $Ge_{1-x}Mn_x$ nanocolumns on Ge(001). *Physical Review B (Condensed Matter and Materials Physics)* 76 (20), 205306 (2007).

82 Park, Y.D. *et al.*, Magnetoresistance of Mn:Ge ferromagnetic nanoclusters in a diluted magnetic semiconductor matrix. *Applied Physics Letters* 78 (18), 2739–2741 (2001).

83 Wang, Y. *et al.*, Direct structural evidences of $Mn_{11}Ge_8$ and Mn_5Ge_2 clusters in $Ge_{0.96}Mn_{0.04}$ thin films. *Applied Physics Letters* 92 (10), 101913 (2008).

84 Pinto, N. *et al.*, Magnetic and electronic transport percolation in epitaxial Ge_{1-x} Mn_x films. *Phys. Rev. B* 72 (16), 165203 (2005).

85 Park, E.S., Kim, D.H., & Kim, W.T., Parameter for glass forming ability of ternary alloy systems. *Applied Physics Letters* 86 (6), 061907 (2005).

86 Lin, H.-T., Huang, W.-J., Wang, S.-H., Lin, H.-H., & Chin, T.-S., Carrier-mediated ferromagnetism in p-Si(100) by sequential ion-implantation of B and Mn. *Journal of Physics: Condensed Matter* 9, 095004 (2008).

87 Verna, A. *et al.*, Ferromagnetism in ion implanted amorphous and nanocrystalline Mn_xGe_{1-x}. *Physical Review B (Condensed Matter and Materials Physics)* 74 (8), 085204 (2006).

88 D'Orazio, F., Lucari, F., Passacantando, M., Santucci, P.P.S., & Verna, A., presented at the Magnetics Conference, 2002 INTERMAG Europe 2002 Digest of Technical Papers. IEEE International, 2002 (unpublished).

89 Liu, L. *et al.*, Investigation of Mn-implanted n-type Ge. *Journal of Crystal Growth* 265 (3–4), 466–470 (2004).

90 Lifeng, L. *et al.*, Magnetic properties of Mn-implanted n-type Ge. *Journal of Crystal Growth* 273 (1–2), 106–110 (2004).

91 Passacantando, M. *et al.*, Growth of ferromagnetic nanoparticles in a diluted magnetic semiconductor obtained by Mn^+ implantation on Ge single crystals. *Physical Review B (Condensed Matter and Materials Physics)* 73 (19), 195207 (2006).

92 Ottaviano, L. *et al.*, Microscopic investigation of the structural and electronic properties of ion implanted Mn–Ge alloys. *Physica Status Solidi (a)* 204 (1), 136–144 (2007).

93 Biegger, E., Staheli, L., Fonin, M., Rudiger, U., & Dedkov, Y.S., Intrinsic ferromagnetism versus phase segregation in Mn-doped Ge. *Journal of Applied Physics* 101 (10), 103912–103915 (2007).

94 Majumdar, S., Das, A.K., & Ray, S.K., Magnetic semiconducting diode of p-$Ge_{1-x}Mn_x$/n-Ge layers on silicon substrate. *Applied Physics Letters* 94 (12), 122505 (2009).

95 Gunnella, R. *et al.*, Magnetization of epitaxial MnGe alloys on Ge(111) substrates. *Surface Science* 577 (1), 22–30 (2005).

96 Liu, L. *et al.*, Growth and properties of magnetron cosputtering grown Mn_xGe_{1-x} on Si(001). *Solid State Communications* 137 (3), 126–128 (2006).

97 Li, H. *et al.*, Magnetic and transport properties of GeMn granular system. *Thin Solid Films* 505 (1–2), 54–56 (2006).

98 Li, H., Wu, Y., Guo, Z., Luo, P., & Wang, S., Magnetic and electrical transport properties of $Ge_{1-x}Mn_x$ thin films. *Journal of Applied Physics* 100 (10), 103908 (2006).

99 D'Orazio, F. *et al.*, Magneto-optical properties of epitaxial Mn_xGe_{1-x} films. *Journal of Magnetism and Magnetic Materials* 262 (1), 158–161 (2003).

100 Zeng, C., Yao, Y., Niu, Q., & Weitering, H.H., Linear magnetization dependence of the intrinsic anomalous Hall effect. *Physical Review Letters* 96 (3), 037204 (2006).

101 Liu, H. & Reinke, P., Formation of manganese nanostructures on the Si(100)-(2 × 1) surface. *Surface Science* 602 (4), 986–992 (2008).

102 Gambardella, P. *et al.*, Surface characterization of Mn_xGe_{1-x} and $Cr_yMn_xGe_{1-x-y}$ dilute magnetic semiconductors. *Physical Review B (Condensed Matter and Materials Physics)* 75 (12), 125211 (2007).

103 Cho, Y.M. *et al.*, Neutron irradiation effects on polycrystalline $Ge_{1-x}Mn_x$ thin films grown by MBE. *Current Applied Physics* 6 (3), 482–485 (2006).

104 Morresi, L. *et al.*, Magnetic and transport polaron percolation in diluted GeMn films. *Materials Science and Engineering: B* 126 (2–3), 197–201 (2006).

105 Sugahara, S., Lee, K.L., Yada, S., & Tanaka, M., Precipitation of amorphous ferromagnetic semiconductor phase in epitaxially grown Mn-doped Ge thin films. *Japanese Journal of Applied Physics Part 2 – Letters Express Letters* 44 (46–49), L1426-L1429 (2005).

106 Wang, Y. *et al.*, Mn behavior in $Ge_{0.96}Mn_{0.04}$ magnetic thin films grown on Si. *Journal of Applied Physics* 103 (6), 066104 (2008).

107 Passacantando, M. *et al.*, Magnetic response of Mn-doped amorphous porous Ge fabricated by ion-implantation. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 257 (1–2), 365–368 (2007).

108 Chen, Y.X. *et al.*, Magnetic and transport properties of homogeneous Mn_xGe_{1-x}

ferromagnetic semiconductor with high Mn concentration. *Applied Physics Letters* 90 (5), 052508 (2007).

109 Verna, A. *et al.*, Magneto-optical investigation of high temperature ion implanted Mn_xGe_{1-x} alloy: evidence for multiple contributions to the magnetic response. *Physica Status Solidi (a)* 204 (1), 145–151 (2007).

110 Pinto, N. *et al.*, Growth and magnetic properties of MnGe films for spintronic application. *Journal of Materials Science: Materials in Electronics* 14 (5), 337–340 (2003).

111 Yu, S.S. *et al.*, Magneto-transport properties of amorphous $Ge_{1-x}Mn_x$ thin films. *Current Applied Physics* 6 (3), 545–548 (2006).

112 Yada, S., Sugahara, S., & Tanaka, M., Magneto-optical and magnetotransport properties of amorphous ferromagnetic semiconductor $Ge_{1-x}Mn_x$ thin films. *Applied Physics Letters* 93 (19), 193108 (2008).

113 Gambardella, P. *et al.*, Paramagnetic Mn impurities on Ge and GaAs surfaces. *Physical Review B* 72 (4), 045337 (2005).

114 Tsui, F. *et al.*, Combinatorial synthesis and characterization of a ternary epitaxial film of Co and Mn doped Ge (001). *Applied Surface Science* 254 (3), 709–713 (2007).

115 Paul, A. & Sanyal, B., Chemical and magnetic interactions in Mn- and Fe-codoped Ge diluted magnetic semiconductors. *Physical Review B (Condensed Matter and Materials Physics)* 79 (21), 214438 (2009).

116 Gareev, R.R. *et al.*, Carrier-induced ferromagnetism in Ge(Mn,Fe) magnetic semiconductor thin-film structures. *Applied Physics Letters* 88 (22), 222508 (2006).

117 Fukushima, T., Sato, K., Katayama-Yoshida, H., & Dederichs, P.H., Spinodal decomposition under layer by layer growth condition and high Curie temperature quasi-one-dimensional nano-structure in dilute magnetic semiconductors. *Japanese Journal of Applied Physics Part 2 – Letters, Express Lett* 45 (12–16), L416-L418 (2006).

118 Holub, M. *et al.*, Mn-doped InAs self-organized diluted magnetic quantum-dot layers with Curie temperatures above 300 K. *Applied Physics Letters* 85 (6), 973–975 (2004).

119 Zheng, Y.H. *et al.*, Cr-doped InAs self-organized diluted magnetic quantum dots with room-temperature ferromagnetism. *Chinese Physics Letters* 24 (7), 2118–2121 (2007).

120 Arrott, A., Criterion for ferromagnetism from observations of magnetic isotherms. *Physical Review* 108 (6), 1394 (1957).

121 Sze, S., *Physics of Semiconductor Devices*, 2nd ed., Wiley, New York (1981).

122 Ovchinnikov, I.V. & Wang, K.L., Voltage sensitivity of Curie temperature in ultrathin metallic films. *Physical Review B (Condensed Matter and Materials Physics)* 80 (1), 012405 (2009).

123 Dietl, T. & Spalek, J., Effect of thermodynamic fluctuations of magnetization on the bound magnetic polaron in dilute magnetic semiconductors. *Physical Review B* 28 (3), 1548 (1983).

124 Kitchen, D. *et al.*, 2007 (unpublished).

125 Lim, S.W., Jeong, M.C., Ham, M.H., & Myoung, J.M., Hole-mediated ferromagnetic properties in $Zn_{1-x}Mn_xO$ thin films. *Japanese Journal of Applied Physics Part 2 – Letters* 43 (2B), L280-L283 (2004).

126 Jeon, H.C. *et al.*, InMnAs diluted magnetic semiconductor quantum dots with

above room temperature ferromagnetic transition. *Advanced Materials* 14 (23), 1725–1728 (2002).

127 Awschalom, D.D., Loss, D., & Samarth, N. (eds), *Semiconductor Spintronics and Quantum Computations*, Springer-Verlag, Berlin (2002).

128 Chiba, D., Matsukura, F., & Ohno, H., Electrical magnetization reversal in ferromagnetic III–V semiconductors. *Journal of Physics D: Applied Physics* 39 (2006).

129 Appelbaum, I. & Monsma, D.J., Transit-time spin field-effect transistor. *Applied Physics Letters* 90 (26), 262501 (2007).

130 Nazmul, A.M., Kobayashi, S., & Tanaka, S.S.M., Electrical and optical control of ferromagnetism in III–V semiconductor heterostructures at high temperature (~100 K). *Japanese Journal of Applied Physics* 43 (2004).

131 Koo, H.C. *et al.*, Control of spin precession in a spin-injected field effect transistor. *Science* 325 (5947), 1515–1518 (2009).

132 Chiba, D. *et al.*, Magnetization vector manipulation by electric fields. *Nature* 455 (7212), 515–518 (2008).

133 Boukari, H. *et al.*, Light and electric field control of ferromagnetism in magnetic quantum structures. *Physical Review Letters* 88 (20), 207204 (2002).

134 Dietl, T., Ohno, H., Matsukura, F., Cibert, J., & Ferrand, D., Zener model description of ferromagnetism in zinc-blende magnetic semiconductors. *Science* 287 (5455), 1019–1022 (2000).

135 Matsukura, F., Ohno, H., Shen, A., & Sugawara, Y., Transport properties and origin of ferromagnetism in (Ga,Mn)As. *Physics Review B* 57 (4), R2037 (1998).

136 Yagi, M., Noba, K.-I., & Kayanuma, Y., Self-consistent theory for ferromagnetism induced by photo-excited carriers. *Journal of Luminescence* 94–95, 523–527 (2001).

137 Dietl, T. & Ohno, H., Engineering magnetism in semiconductors. *Materials Today* 9 (11), 18–26 (2006).

138 Weisheit, M. *et al.*, Electric field-induced modification of magnetism in thin-film ferromagnets. *Science* 315 (5810), 349–351 (2007).

Index

absolute energy scale, 30
absolute supersaturation, 47
activation enthalpy, 309, 328–9
adatoms, 52, 57, 151, 158
adsorbed atoms *see* adatoms
adsorption equilibrium constants, 127
advanced chemical vapour deposition methods, 8
advanced CMOS technologies, 117
Al-induced crystallisation process, 200
alloy limit, 275
APM cleaning, 153, 155
Arrhenius expression, 301
Arrhenius plots, 194
  nucleation incubation times, 195
Arrott plots, 586, 592
Asaro–Tiller–Grinfeld instability, 66, 253
atmospheric pressure chemical vapour deposition, 118
atomic force microscopy, 594
atomistic Monte Carlo simulations, 232

Bales–Zangwill instability, 62
Bales–Zangwill theory, 62
band gap, 404
Barton, Cabrera, and Frank model, 56–7
B–B adsorption, 137
biaxial compression, 347
biaxial tension, 347, 348
BiCMOS *see* bipolar complementary metal oxide semiconductors
bipolar complementary metal oxide semiconductors, 118, 123, 124, 486–90
  device characteristics, 489–90
  device structure, 486–7
    self-aligned SEG SiGe HBT, CMOS transistor, and poly-Si resistor, 487
  fabrication process, 487–9
    integration process schemes, 488
Bir–Pikus shear deformation potentials, 35
Boltzmann constant, 47
boron, 327
BOX layers *see* buried oxide layers
Bragg reflector, 555
Bragg's law, 19
Bridgman's crystal growth technique, 582
Burgers vector, 9, 11, 12, 14
buried oxide layers, 179

capillarity model, 213
carbon, 476–7
carrier devices, 339
cathodoluminescence imaging, 340
chemical dry etching, 178
chemical mechanical planarisation, 153
chemical–mechanical polishing method, 185
chemical vapour deposition
  atomic layer doping, 135–8
    B dose in SiGe, 138
    B peaks SIMS profile in SiGe, 138
    P peaks SIMS profile in SiGE for PH_3 exposure temperature, 139
    process sequence using batch reactor, 136
    surface reaction-limited processing, 136
  B_2H_6 and PH_3 partial pressures
    deposition rate, 129

        dopant concentration in $Si_{1-x}Ge_x$ films, 131
        Ge fraction, 130
    doping, 128–39
    epitaxial growth techniques, 119–21
        chemical equilibrium data for $Si/H_2O/SiO_2$ system, 120
        different techniques deposition and base pressure, 121
    future trends, 139–40
    GeH_4 partial pressure
        Ge fraction, 126
        SiGe growth rate, 126
    HBT base doping and dopant profile control, 130–5
        B concentration comparison in SiGe (20% Ge), 133
    heteroepitaxy, 122–8
        epitaxial growth rate for Si and SiGe, 125
        low-temperature SiGe growth, 124–8
        Si(100), Si(110), Si(110) SiGe growth rate, 128
        SiGe epitaxial deposition process sequence, 122
        silicon surface preparation, 123–4
        standard cleaning procedures contaminant removal and chemistry, 123
    in-situ doping, 128–30, 131, 132
        B and P carrier concentrations in $Si_{1-x}Ge_x$ films, 132
    main steps, 119
    silicon–germanium crystal growth, 117–40
        HBT integrated TEM cross-section in BiCMOS technology, 118
    SIMS profile
        B in SiGe and SiGe:C for two C concentrations, 134
        P in SiGe and SiGe:C for two C concentrations, 134
CMOS *see* complementary metal oxide semiconductor
CMOS technology, 559
CMP *see* chemical mechanical planarisation
CMP method *see* chemical–mechanical polishing method
cobalt disilicide, 460
complementary BiCMOS (CBiCMOS), 130–1
complementary metal oxide semiconductor, 119, 121, 123, 124, 166, 190, 425, 433, 499–524
compliant effect, 159, 165
compositionally graded buffer, 149
conduction-band minima, 27
conduction electron spin resonance, 408
conductivity mass, 38
continuous-epitaxial-growth, 477–9
$COSi_2$ *see* cobalt disilicide
Coulomb scattering, 512, 543
critical nucleus, 49, 60
critical thickness, 8–17
    compressive film strain vs film thickness, 15
    cubic group IV elements and compounds, 9
    dislocation half-loops and movement nucleation to the interface, 16
    edge dislocation formed by inserting an extra half-layer of atoms, 9
    under equilibrium vs obtained at low growth temperatures, 16
    right-handed screw dislocation formed by a crystal displacement, 11
    Thompson's reference tetrahedron, 10
crosshatch, 151–5
    pattern, 151
crystal growth mechanisms
    adsorption and desorption, 51–4
        energetics, 52
        Si film growth rate temperature dependence, 53
    epitaxial growth kinetics, 60–4
        Ehrlich–Schwoebel barriers, 61
    fundamental growth processes, 50–60
        nucleation, 59–60
    heteroepitaxy, 64–8
        growth strain relaxation mechanisms, 66
    surface diffusion and step motion, 56–9
        adatom potential energy 1-D representation, 57

  - surface structure and fundamental growth modes, 54–6
    - adsorption positions types on the surface, 54
    - epitaxial growth modes, 55
  - thermodynamics, 46–50
    - film growth basic modes, 50
    - Gibbs free energy change, 49
  - understanding in silicon–germanium nanostructures, 45–68
- crystal structure, 4
  - diamond structure, 4
- Curie temperature, 206, 207, 576
- CVD *see* chemical vapour deposition
- Czochralski method, 74–5
  - illustration, 75

- degree of strain relaxation, 101
- density functional theory, 211, 579
- density of states, 580
- density-of-states masses, 40
- DFT *see* density functional theory
- diffusion mechanism, 300–2
  - dopant diffusion mechanisms, 300
  - self-diffusion mechanisms, 300
- diluted HF, 124
- dilute magnetic semiconductor
  - spintronics of nanostructured manganese, 575–602
    - electric field-controlled ferromagnetism, 596–601
    - ferromagnetism theories, 578–81
    - future trends, 601–2
    - growth and characterisations, 581–96
    - MnGe growth and characterisation, 584
- Dingle ratio, 418
- Dingle temperature, 366
- directional solidification, 78
- disorder-induced gap states, 466
- domes, 92
- dominant scattering mechanism, 376
- dopant activation, 322–3
- dopant diffusion, 299–32
  - germanium, 316–23
    - active As concentration profiles, 317
    - dopant activation, 322–3
    - epitaxially grown Ge, 319
    - group III atoms, 319–20
    - group V atoms, 317–19
    - implanted B SIMS profiles, 321
    - P, As, and Sn intrinsic diffusivities, 318
    - P, As, Sb and B in Ge intrinsic diffusivities, 319
    - P concentration profiles, 323
    - SIMS depth profiles, 320
    - transient enhanced diffusion, 320–2
  - silicon–germanium alloys, 323–9, 330
    - activation enthalpies, 329
    - activation enthalpy of diffusion, 309, 328–9
    - Arrhenius plot of intrinsic As diffusivities, 325
    - B concentration in strained multilayer structure with different Ge content, 327
    - B concentration profiles in an epitaxial Si layer, 330
    - Ge composition dependence, 306–8
    - group III atoms, 326–8
    - group V atoms, 323–6
    - intrinsic B diffusivity, 328
    - Sb profiles, 324
    - time dependence, 326
- dopant segregation, 329–30
- doped rare-earth ions, 566
- DOS *see* density-of-states masses
- Drude model, 401
- dual plateaus with an intermediately graded slope, 482
- duMond–Hart–Bartels, 17

- EBIC *see* electron beam induced current
- Ehrlich–Schwoebel barrier, 61
- Ehrlich–Schwoebel instability, 61
- EIES *see* electron impact emission spectroscopy
- elasticity theory, 240
- elastic relaxation, 215
- elastic theory, 30
- elective-epitaxial-growth, 481
- electric field-assisted MILC, 198
- electric field-controlled ferromagnetism, 578, 596–601
  - Mn-implanted Ge dots, 597–9

    magnetic moment vs applied field, 598
    magnetic properties, 598
  self-assembled DMS MnGe QDs, 599–600
    ferromagnetism control, 600
electro-absorption, 558
electroluminescence, 447–9
electron backscattering diffraction, 200
electron beam induced current, 340
electron gun evaporator, 84
electronic band structures
  band structures, 26–9
    bulk silicon and germanium, 26–8
    SiGe alloys, 28–9
    $Si_{1-x}Ge_x$ alloys low-temperature bandgap vs composition, 29
  effective mass, 38–41
    hole DOS masses Ge fraction dependence, 40
  Si and Ge
    empirical tight-binding method, 27
    many-valley structure, 27
  silicon–germanium alloys, 26–41
  strained $Si_{1-x}Ge_x$ on Si substrate, 30–5
    band-edge levels, 35
    band lineups in a strained SiGe layer on a Si substrate, 33
    band lineups in a strained Si layer on a SiGe substrate, 33
    conductivity masses Ge fraction dependence, 39
    film lattice constant, 30
    layer lattice constants and strain tensor components, 32
    strained Si layer band lineups, 33
  strain effects, 29–37
    band alignments for biaxial compressively and tensile strained on $Si_{1-x}Ge_x$ layer, 37
    band offsets, 36–7
    (100)-strained $Si_{1-x}Ge_x$ layer fundamental bandgap in eV, 36
    strained $Si_{1-x}Ge_x$ on $Si_{1-y}Ge_y$ substrate, 36
electronic-toll-collection, 491
electron impact emission spectroscopy, 86
electro-optic modulators, 556
energy dispersive X-ray spectroscopy, 591
energy shift, 31
enhanced Secco etching, 180
epilayer, 250
epitaxial growth, 474–8
  carbon-doped SiGe growth, 476–7
    effective diffusivities, 477
  continuous growth of Si/SiGe multilayer, 477–9
    dopant concentrations and Ge content in a SiGe HBT, 478
  SiGe techniques, 474–6
    $Si_{0.85}Ge_{0.15}$ grown on Si(100) substrate, 475
  techniques
    atmospheric pressure CVD, 119–22
    low pressure CVD, 119–22
    plasma activated processing, 119–22
    ultra high vacuum, 119–22
Eshelby's formalism, 233
etch-pit-density (EPD), 180, 184
extended X-ray absorption fine-structure, 6

Fabry-Pérot filter, 555
faceted islands, 88
Fang–Howard wave function, 420
FEM *see* finite element methods
FEM solver, 234
Fermi level effect, 301
Fermi level pinning, 465, 466
Fermi wave numbers, 367
ferromagnetism, 578–81
FET *see* field effect transistors
Fick's laws, 313
field cooled magnetisation's, 583, 592
field effect transistors, 191, 339, 424, 499–524, 528, 596–7
figure of merit, 484, 485
finite-difference-time-domain, 444
finite element methods, 211, 258
float zone method, 75–6
  illustration, 76
FLP *see* Fermi level pinning
flux control, 86–7
FOM *see* figure of merit
Fourier transform, 365
fractional quantum hall effect (FQHE), 410
Frank–van der Merwe, 251

Frank–van-der-Merwe growth, 46, 50
Franz–Keldysh effect, 558
Friedel oscillations, 368
full-potential linearized augmented plane wave, 578
full-width-half-maximum (FWHM), 442

gas-source method, 85–6
gas-source molecular beam epitaxy, 46, 52–4
Gatan precision ion polishing system, 585
gate stack formation, 529–36
Ge condensation method, 172
  SGOI formation, 172–86
    fundamentals, 172–5
germanide, 456
  ultra-large-scale integrated (ULSI) applications, 456–66
germanium
  MOSFET, 528–49
    and electron inversion layer mobility, 536–43
    future trends, 548–9
    gate stack information, 529–36
    Schottky interface, 544–8
  self-diffusion and dopant diffusion, 299–32
    diffusion mechanism, 300–2
    dopant diffusion in germanium, 316–23
    dopant diffusion in silicon–germanium (SiGe) alloys, 323–9
    dopant segregation, 329–30
    interdiffusion studies, 311
    self-diffusion in germanium, 302–6
    self-diffusion in silicon–germanium (SiGe) alloys, 306–9
    silicon–germanium interdiffusion, 310–16
germanium dots, 260–3, 434–45
germanium island
  epitaxial growth modelling consideration, 211–13
  future trends: ordering Ge islands on pit-patterned Si(001), 241–3
    islands AFM 3D topography, 242
    shaded regions energy per atoms, 243
  silicon(001) thin films, 211–43
  Stranski–Krastanow heteroepitaxy, 213–35
    beyond onset: SiGe intermixing, 232–5
    beyond onset: vertical and horizontal ordering for application, 235–41
    onset, 219–32
germanium on insulator, 186, 199
germanosilicide
  thin films formation, 457–8
Gibbs free energy, 47, 59, 60
Gibbs–Thomson equation, 48
GISAXS *see* grazing-incidence small-angle X-ray scattering
GIXD, grazing–incidence X–ray diffraction
global modelling, 74
graded buffer method, 158
grazing-incidence small-angle X-ray scattering, 258
grazing-incidence X-ray diffraction, 258
growth methods
  application to heteroepitaxy, 79
  Czochralski method, 74–5
    illustration, 75
  float zone method, 75–6
    illustration, 76
  multicomponent zone melting method, 76–9
    fundamental concept, 77
    Ge content distribution, 77
    orientation distribution comparison, 78
  SiGe phase diagram, 73
  types and applications of silicon-germanium bulk crystal, 72–9

Hall effect, 581
Hall effect electron mobility, 543
HBT *see* heterojunction bipolar transistor
heavy-hole band, 28
heterobipolar transistor, 23, 424
heteroepitaxy, 64–8, 79, 250–2
  silicon–germanium, by CVD, 122–8
    epitaxial growth rate for Si and SiGe, 125
    low-temperature SiGe growth, 124–8

    - Si(100), Si(110), Si(110) SiGe growth rate, 128
    - SiGe epitaxial deposition process sequence, 122
    - silicon surface preparation, 123–4
    - standard cleaning procedures contaminant removal and chemistry, 123
  - Stranski-Krastanow, 213–35
- heterojunction bipolar transistor, 117, 118, 479–86
  - device structure, 479, 480
    - self-aligned SiGe HBTs cross-sections, 480
  - fabrication process, 479–82
    - SiGe HBTs with self-aligned structures, 481
  - transistor characteristics, 482–6
    - cutoff frequency and maximum oscillation frequency, 483
    - cutoff frequency on collector–emitter breakdown voltage, 484
    - SiGe and SiGeC HBTs intrinsic impurity profiles, 482
- Heusler alloy, 206–7
- HH band *see* heavy-hole band
- high resolution X-ray diffraction, 17
- homoepitaxial layers, 250–1
- Hooke's law, 18
- hut clusters, 88
- hydrogen desorption, 125

- IC *see* integrated circuit
- impurity scattering, 362
- infrared sensors, 563
- *in-situ* scanning tunnelling microscopy, 84
- integer quantum Hall effect (IQHE), 410
- integrated circuit
  - SiGe HBT and BiCMOS, 490–3
    - optical-fibre links, 490–1
    - transmitter and receiver built with the IC chipset, 492
    - wireless communications, 491–3
- interdiffusion
  - silicon–germanium, 310–16
    - Ge composition dependence, 310–13
    - Ge concentration profiles at 750°C, 313
    - Ge concentration profiles at 1030°C, 313
    - Ge profiles of heterostructures, 315
    - Ge self-diffusivities Arrhenius plot, 316
    - interdiffusivities from strained SLs, 312
    - strain effect, 314–16
    - structure and strain distribution, 312
    - studies, 312
- interface behaviour, 466
- interface roughness, 543
- interface-roughness scattering, 362
- interfacial energy per unit area, 49
- International Technology Roadmap for Semiconductors (ITRS), 483
- ion beam irradiation techniques, 193
- ion-implantation buffer, 158, 161–5
- ion-implantation defects, 161
- ion-implantation methods, 164, 166
- IRS *see* interface-roughness scattering
- IS *see* impurity scattering
- island engineering, 260–8
- island evolution, 252–60
- isotope engineering, 302–3

- junction leakage, 463
- JV-HRXRD analysis software, 19

- kinetic Monte Carlo simulations, 231
- Knudsen cell (K-cell), 84

- Landau level index, 412
- Landau splitting, 411
- Langmuir-type adsorption scheme, 127
- large-scale integrated circuits, 552
- large-scale integration
  - SiGe HBT and BiCMOS, 490–3
    - optical-fibre links, 490–1
    - wireless communications, 491–3
- laterally diffused MOS, 489
- lateral ordering, 263–8
- lattice parameters, 4–7
  - elastic stiffness constant, 7
  - Si1–*x*Ge*x* alloys at 25°C, 5
- LEED *see* low energy electron diffraction
- LEPTOS, 20
- LH band *see* light-hole band
- light-emitting devices, 443

light-hole band, 28
light sources, 565–6
  doping rare earth, 566
  Si light sources, 565–6
linear elasticity theory, 11
line defects, 339
line-splitting, 262
liquid-phase crystallisation (LPC), 192
liquid phase diffusion, 78
local spin-density approximation, 580, 582
local strain technology, 504–5
Lorentz model, 371
low energy electron diffraction, 254
low-energy plasma-enhanced chemical vapor deposition, 422
low pressure chemical vapour deposition, 118
low-temperature buffer, 158–61
LSI *see* large-scale integration
Luttinger parameters, 34

Mach–Zehnder interferometer, 557
magnetic force microscopy, 594
magnetisation, 576
magneto-resistance ratio, 381
mass spectrometer, 86
Mathiessen's rule, 401
MBE *see* molecular beam epitaxy
metal-imprint techniques, 193
metal-induced gap states, 466
metal-induced lateral crystallisation, 192, 194–6
metal-insulator semiconductor field-effect transistor, 272
metal-insulator transition, 362, 369, 371–2, 381–3
metal oxide semiconductor field effect transistors, 94, 98, 122, 127, 171, 310, 456, 465, 500–1
  fabrication and electron inversion layer mobility, 536–43
    Au/GeO$_2$/Ge(100) MOS capacitors bi-directional C-characteristics, 537
    conduction-band energy minima and equi-energy surface, 541
    Ge and Si surface roughness, 543
    Ge(100) n-MOSFET extracted electron mobility, 540
    Ge n-MOSFETs process flow, 538
    Ge(100) n-MOSFETs transfer characteristics, 539
    substrate orientation effect on electron mobility, 541
    surface oxide thickness on Ge(100), 542
  gate stack formation, 529–36
    Au/GeO$_2$/Ge MOS capacitor characteristics, 536
    Au/GeO$_2$/Ge(100) MOS capacitor C–V characteristics, 535
    extinction coefficients, 535
    Ge/GeO$_2$ stack for ^{18}O$_2$ oxidation, 530
    Ge/GeO$_2$ system thermodynamic consideration, 532–6
    GeO$_2$/Ge sample with oxygen isotope, 532
    GeO vapor pressure in equilibrium with GeO2, 534
    Ge oxidation and GeO desorption, 529–32
    line-patterned GeO$_2$ film on Ge, 533
    thermal desorption spectroscopy, 531
  germanium metal Schottky interface and metal source/drain, 544–48
    Fermi-level pinning at metal/Si and metal/Ge interfaces, 545
    FLP position, 545
    I_S–V_{DS} characteristics, 548
    metal/Ge interface, 544–7
    metal source/drain Ge p- and n-MOSFETs, 547
    metal source/drain MOSFETs, 547–8
    ultra-thin oxide thickness dependence of *J*–*V* characteristics, 546
  high electron mobility germanium, 528–49
    future trends, 548–9
micro-indentation techniques, 193
micro-origami, 269
micro Raman spectroscopy, 155
microscopic Raman spectra analysis, 79
MILC *see* metal-induced lateral crystallisation

Miller indexes, 216
misfit dislocations, 13, 14, 17, 181, 351–2
MIT *see* metal-insulator transition
Mn-doped Ge films, 585–90
  fast Fourier transformation map, 586
  HR-TEM images, 589
  SQUID measurement, 588, 590
  TEM cross section of as-grown and 400°C, 587
$Mn_{0.05}Ge_{0.95}$ quantum dots, 590–5
  AFM and MFM images, 595
  magnetic properties, 593
  structural properties, 591
model-solid theory, 30
modified Frank Read (MFR), 150
modulation doped quantum well structures, 400
modulation-doping, 407
modulators, 556–9
  electro-absorption, 558
  free carrier absorption, 556–7
  other high-speed electro-optic materials, 558–9
  resonant structure, 557–8
  thermo-optical effect, 557
molecular beam epitaxy, 8, 202, 581
  clean Ge surface STM image
    oxygen exposure of 6 L, 111
    oxygen reaction modes, 112
    structure, domain boundary and RHEED pattern, 109
    subjected to oxygen exposure of 1.2 L, 109
  future trends, 113
  germanium growth on silicon, 87–94
    plan-view bright-field TEM image, 91
  GeSn growth on Si for strained Ge, 98–104
    GeSn layers with diffraction vector TEM images, 99
    strained-Ge morphology and interfaces image, 103
  macro-island
    cross-sectional HRTEM image, 91
    formation mechanism in Ge/Si(001) system, 93
  7-ML Ge sample
    cross-sectional HRTEM image, 90
  hot clusters and extra spots TEM/TED image, 89
  plan-view bright-fieldTEM image grown at 600ºC, 92
MnGe, 584–95
  self-assembled DMS MnGe quantum dots, 590–5
  thin and thick films, 585–90
nanoscale textures in SiGe/Si systems, 94–8
  SiGe layer measured tilt angles, 97
  SiGe/Si(001) interface plan-view TEM images, 97
  XRMD-2DRSM around Si 004 and SiGe 004 diffraction spots, 96
nanostructured formation, 87–113
  germanium growth on silicon, 87–94
  GeSn growth on Si, 98–104
  oxygen reaction with clean Ge surfaces, 108–13
  SiGeC growth on Si, 104–8
  SiGe/Si system nanoscale textures, 94–8
oxygen reaction with clean Ge surfaces, 108–13
  dual-bias STM images, 110
  SiGeC growth on Si, 104–8
  silicon–germanium crystal growth, 83–113
  techniques, 84–7
SiGeC growth on Si
  5-ML-thick $Si_{0.769}Ge_{0.183}C_{0.048}$ layers, 107
  sample with a $Si_{0.473}Ge_{0.473}C_{0.054}$ (6 ML)/ $Si_{0.5}Ge_{0.5}$(1 ML)/Si(001) structure image, 107
  $Si_{0.5}Ge_{0.5}$ and $Si_{0.489}Ge_{0.489}C_{0.022}$ films surface structures images, 105
  $Si_{0.478}Ge_{0.478}C_{0.044}$ films STM image, 106
single $Ge_{1-x}Sn_x$ layers
  dislocation morphology and interfaces TEM images, 103
  $Ge_{1-x}Sn_x$ 224 reciprocal lattice points with constant Sn content on v-Ge, 100
  $Ge_{1-x}Sn_x$ 224 reciprocal lattice

        points with Sn content variation and strain relaxation behaviour, 102
      XRD-2DRSM for 224 reflections, 103
  techniques
    flux control and monitoring, 86–7
    gas-source method, 85–6
    single complete layer formation, 87
    solid-source method, 84–5
molecular dynamics, 211
Monte Carlo simulations, 5
Moore's law, 45, 400
MOSFET *see* metal oxide semiconductor field effect transistors
mounds, 213
multicomponent zone melting method, 76–9
  fundamental concept, 77
  Ge content distribution, 77
  orientation distribution comparison, 78
multigate structures, 501
multiple-scattering effects, 369–71
M x N reconstruction, 217

nanotechnology
  rolled-up, 268–72
narrow line effect, 459
nickel monosilicide, 460–2
NiSi *see* nickel monosilicide
Nomarski optical micrographs, 196

opt-electronic integrated circuits (OEIC), 148
optical devices
  silicon and germanium, 551–67
    future trends, 566–7
    light sources, 565–6
    modulators, 556–9
    optical waveguides, 552–6
    photodetectors and photovoltaics, 559–65
optical-fibre links, 490–1
  parameters of passive elements based on SiGe BiCMOS, 491
optical microcavity, 434
optical waveguides, 449–51, 552–6
  Si and SiON waveguide, 554
  waveguide structures, 553–4
    rectangular waveguides, 553
    ridge, rectangular and slotted, 553
    ridge waveguides, 553
    slab waveguides, 553
    wavelength and sub-wavelength-scale structures, 553
  wavelength and sub-wavelength structures, 555–6
order of desorption, 52

pair diffusion model, 301
palladium, 462
p-channel MOSFET, 160
percolation theory, 579
phase diagram, 7–8
  liquidus–solidus curve, 7
Philips simulation software, 6
photodetectors/photovoltaics, 559–65
  Ge photodiodes, 562–3
    band gap structures of Ge and strained n-type Ge, 562
  infrared sensors, 563
  Si photodiodes, 559–62
    light confinement by resonators, 560
    structures of light confinement, 561
  solar cells, 563–5
    plasmon solar cell structure, 564
photoluminescence (PL) spectra, 219
photonic crystal microcavities, 440–3
photonic crystals, 555
planar defects, 339
planarisation effect, 155
platinum, 461, 462
Poisson ratio, 215
polycrystallisation, 73, 74, 78
post-CMP cleaning, 153
post-deposition annealing, 99
power consumption, 499
power supply voltage, 499
pre-ion-implantation method, 164
prepyramids, 213, 252
pseudomorphic layer, 29
Purcell effect, 437
pyrolytic boron nitride, 84

QDSLs *see* quantum dot superlattices
quantum cascade laser structures, 445–51
quantum computing, 399
quantum-confined Stark effect, 558

quantum dots, 351–2, 355–6, 577
quantum dot superlattices, 352–4
quantum Hall ferromagnetism (QHF), 414
quantum point contacts, 415–17
quantum wells, 249, 361, 372–9
quartz crystal microbalance, 86
Quartz temperature, 86
quasi-ballistic transport, 500
QW *see* quantum wells

radial hybrid superlattice (RSL), 270
Raman mapping, 152, 155
Raman spectroscopy, 177
random alloy scattering, 420
random-phase approximation (RPA), 364
random potential, 365
rapid thermal annealing (RTA), 121
rapid thermal CVD technique, 121
rate constant of desorption, 52
REBOL A, 271
reduced pressure chemical vapour deposition, 118
reduced surface field, 485
reflection high energy electron diffraction, 86, 202, 263
refractive index, 557
relaxation ratio, 176
RHEED *see* reflection high energy electron diffraction
Ruderman–Kittel–Kasuya–Yosida interaction and percolation theory, 578
Rutherford backscattering, 5

sacrificial-emitter process, 479, 480, 481
salicide, 457
  formation in MOSFETs, 457
SBH *see* Schottky barrier height
scanning electron microscopy (SEM), 197
scanning tunneling microscopy, 88, 254–5
scattering times, 366–7
Schottky barrier height, 462–6
Schottky depletion layer., 560
SEG *see* selective epitaxial growth
SEG-SA, 479
selective epitaxial growth, 474, 475, 479, 480
self-aligned silicide *see* salicide
self-diffusion, 299–32
  germanium, 302–6
    charge states of vacancies, 305–6
    Ge SL structures, 303
    intrinsic Ge self-diffusivities, 304
    intrinsic Ge self-diffusivity, 303–5
    intrinsic self-diffusivity, 304
    isotope engineering, 302–3
    SIMS and simulated depth profiles, 306
  silicon–germanium alloys, 306–9
    composition dependence of activation enthalpies, 309
    self-diffusivities, 308
    SIMS concentration profiles, 307
    temperature dependence, 308
SEP *see* sacrificial-emitter process
SEP-SA, 479, 480
SGOI *see* SiGe on insulator
shape anisotropy, 205
shear stress, 12
Shockley partial, 347, 348, 350
Shubnikov–de Haas oscillations, 366
SiGe *see* silicon–germanium
SiGe dot-based FET, 273
SiGe FinFET, 517–20
SiGe nanowire MOSFETs, 520–4
SiGe on insulator, 172, 511–17
SiGe p-channel MOSFETs, 506–11
SiGe/Si heteroepitaxy, 259
silicides
  crystalline and electrical properties, 459
  crystalline properties, 459–62
    $CoSi_2$, 460
    NiSi, 460–2
    Pd, Pt, Ir, and Er silicides, 465
    $TiSi_2$, 459–60
  thin films formation, 457–8
  ultra-large-scale integrated (ULSI) applications, 456–66
silicon–germanium, 122
  as channel materials and stressors in MOSFETs, 502–6
    calculated phonon-limited 300 K on (001) Ge substrates, 505
    calculated phonon-limited 300 K on (001) Si substrates, 504
    channel material, 502
    strained-Si, 506
    stressor, 503–6

  - FET and CMOS technologies, 499–524
    - calculated phonon-limited 300K, 503
    - necessity for high mobility channels, 499–501
  - HBT and BiCMOS technologies, 473–94
    - bipolar complementary metal oxide semiconductors, 486–90
    - epitaxial growth, 474–8
    - heterojunction bipolar transistor, 479–86
    - IC and LSI applications, 490–3
  - (SiGe) channel MOSFETs, 506–24
    - average injection velocity vs extracted hole velocities, 510
    - channel metal S/D p-MOSFET, 514
    - 3D multigate CMOS structure, 519
    - dual hetero-channel CMOS structure, 512
    - effective hole mobility, 508
    - experimental hole mobility enhancement factor, 518
    - extracted hole mobilities, 508
    - fabricated SiGe FinFET, 519
    - fabrication flow of wire-shaped SiGe MOSFETs, 522
    - fabricaton flow, 515
    - G_m–V_g characteristics, 522
    - lateral strain relaxation, 516
    - measured trans-conductance (G_m) values, 521
    - SGOI channel with Ge fraction, 513
    - SGOI p-MOSFETs effective hole mobility, 513
    - SGOI p-MOSFETs mobility, 515
    - SiGe FinFET, tri-gate MOSFETs, 517–20
    - SiGe MOS structures, 507
    - SiGe nanowire MOSFETs, 520–4
    - SiGe n-channel MOSFETs, 511
    - SiGe–NWs after thermal oxidation, 523
    - SiGe-on-insulator MOSFETs, 511–17
    - SiGe p-channel MOSFETs, 506–11
- silicon–germanium alloys, 323–9
  - electronic band structures, 26–41
    - band structures, 26–9
    - effective mass, 38–41
    - strain effects, 29–38
  - self-diffusion and dopant diffusion, 299–32
    - diffusion mechanism, 300–2
    - dopant diffusion in germanium, 316–23
    - dopant diffusion in silicon–germanium (SiGe) alloys, 323–9
    - dopant segregation, 329–30
    - self-diffusion in germanium, 302–6
    - self-diffusion in silicon–germanium (SiGe) alloys, 306–9
    - silicon–germanium interdiffusion, 310–16
- silicon–germanium bulk crystal
  - growth methods types and applications, 72–80
    - application to heteroepitaxy, 79
    - growth methods, 74–9
    - phase diagram, 73
- silicon–germanium crystal
  - chemical vapour deposition, 117–40
    - epitaxial growth techniques, 119–21
    - future trends, 139–40
    - HBT integrated TEM cross-section in BiCMOS technology, 118
    - silicon–germanium doping, 128–39
    - silicon–germanium heteroepitaxy, 122–8
  - ferromagnetic Heusler alloys heteroepitaxial growth for (SiGe)-based spintronic devices, 201–7
    - χ_{min} values obtained from stoichiometric samples, 203
    - $Fe_2MnSi/Ge(111)$ image and temperature dependence, 207
    - Fe_3S atomically controlled epitaxial growth on Ge substrates, 202–4
    - Fe_3Si/Ge, Si heterostructures magnetic and electrical properties, 204–6
    - $Fe_3Si/Ge(111)$ TEM image and Fe and GE concentration profiles, 204
    - Fe_3Si layer magnetic hysteresis loop and crystal quality dependent coercivities, 205

    - Fe_3Si/Si Schottky contact image and non-local spin signal, 206
    - full-Heusler alloys atomically controlled epitaxia growth, 206–7
    - spin transistor structure, 202
  - miscellaneous methods and materials on thin films, 190–208
  - molecular beam epitaxy, 83–113
    - future trends, 113
    - nanostructured formation, 87–113
    - techniques, 84–7
  - oriented growth for thin film transistors and 3-D stacked devices, 191–201
    - a-SiGe SPC temperature on SiO_2, 192
    - electric-field assisted milc for uniform and nanowire growth, 196–8
    - enhanced crystal nucleation for low-temperature solid-phase crystallisation, 192–4
    - hole concentration and mobility from seeding edge, 201
    - incubation times Arrhenius plots for nucleation in a-SiGe, 195
    - metal-induced lateral crystallisation for SiGe with large grain, 194–6
    - Nomarski optical micrographs and sample structure, 197
    - sample structure and lateral growth velocity at the cathode, 198
    - sample structure cross-section and SiGe optical micrograph, 196
    - sample structures cross-section, 193
    - SiGe-mixing-triggered rapid melting growth for defect-free Ge on insulator, 199–201
    - ultra-high-speed Ge channel devices progress, 191
    - X-TEM image and depth distribution, 199
- silicon–germanium layers
  - growth on thin silicon on insulator substrate, 171–88
    - demand for virtual substrate and (Si)Ge on insulator, 171–2
    - extension toward Ge on insulator, 186–7
    - SGOI formation by Ge condensation method, 172–86
- silicon–germanium microstructures
  - strain engineering, 247–82
    - growth insights, 250–60
    - island engineering, 260–8
    - potential applications, 272–9
    - related information, 279–82
    - rolled-up nanotechnology, 268–72
- silicon–germanium nanostructures
  - carrier transport in silicon/silicon–germanium devices, 424–7
    - heterobipolar transistors, 424
    - modulation-doped Si/SiGe field effect transistors, 424–5
    - strained CMOS, 425–7
    - stressor techniques, 426
  - comparison with experimental results, 383–9
    - characteristic mobility and critical density, 386
    - discussion of the metal–insulator transition, 388–9
    - electron densities vs impurity density, 384
    - experiments with heterostructures, 387–8
    - experiments with quantum wells, 383–7
    - peak mobility vs peak electron density, 383
    - Si/SiGe QWs and HSs characteristic mobility and critical density, 386
    - ZnO/MgZnO HSs characteristic mobility and critical density, 391
  - compressive strain vs tensile strain, 345
    - AFM images of $Si_{0.5}Ge_{0.5}$, 345
  - crystal growth mechanisms, 45–68
    - epitaxial growth kinetics, 60–4
    - fundamental growth processes, 50–60
    - growth thermodynamics, 46–50
    - heteroepitaxy, 64–8
  - dislocations and other strain-induced defects, 338–56
    - extended defects on SiGe devices, 339–41

- future trends, 354–6
- historical overview, 341–6
- line defects, 339
- planar defects, 339

feasibility of attaining high-quality relaxed SiGe on Si(111) and Si(110), 349–51
- PVTEM of $Si_{0.97}Ge_{0.03}$ on Si(111), 350
- Schimmel-etched $Si_{0.97}Ge_{0.03}$, 350

future trends, 354–6
- coherent islands formed on Ge(111), 356
- heterostructure nanowires defects, 354–5
- self-assembled Si quantum dots, 355–6
- Thompson tetrahedron for visualisation of shear stresses, 355

germanium dots microcavity photonic devices, 434–45
- current-injected devices, 443–5
- fabricated microdisk resonator with Ge quantum dots, 437
- fabricated microring resonator with Ge dots, 439
- Ge dots in DBR cavities, 435–6
- Ge dots microdisk EL device, 444–5
- Ge self-assembled quantum dots, 435
- microdisk and microring resonators, 437–40
- μ-PL spectra from T6 microcavities, 443
- μ-PL spectra of a 4-μm microdisk with Ge dots, 438
- μ-PL spectra of a 3-μm microring resonator, 440
- μ-PL spectrum from T6 microcavity, 442
- PhC microcavity, 441
- photonic crystal microcavities, 440–3
- planar cavity with Ge dots, 436

heterostructure transport, 379–83
- lowest-order results, 380–1
- metal–insulator transition, 381–3
- resistance ratio, 382
- scattering time ratio, 382

historical overview, 341–6
- historical observations, 346

low-dimensional transport, 407–23
- conductance, 416
- conduction band variation, 408
- diagonal resistivity and Hall resistivity, 411
- effective hole mass, 423
- Hall mobilities temperature variation, 422
- hole mobilities vs Ge content, 419
- measured gap, 412
- two-dimensional electron gases, 407–19
- two-dimensional hole gases, 419–23
- valley splitting density dependence, 413

low-TDD compositionally graded buffers, 344
- SiGe/Si(001), 344

metastable nature of SiGe layers on Si(001), 341–2
- growth temperature on relaxation, 342

microcavities and quantum cascade laser structures, 433–52

misfit dislocations as a template for quantum dot growth, 351–2
- sample with 1.0 nm of Ge coverage grown at 750°C, 351

mobility vs electron density, 373
- different expressions for the local-field correction, 373
- different positions of the impurity layer, 374
- heterostructure for IS, 380
- heterostructure with IS and IRS, 381
- IRS are $\Delta = 3$ Å and $\Lambda = 60$ Å, 374
- $\Lambda = 60$ Å, $\Lambda = 30$ Å, and $\Lambda = 15$ Å, 375

model, disorder and transport theory, 363–72
- disorder, 365–6
- metal-insulator transition, 371–2
- model, 363–5
- multiple-scattering effects, 369–71
- scattering times, 366–7
- spin-polarised electron gas, 367–8

- temperature dependence of the mobility, 368–9
quantum cascade laser structures, 445–51
- conduction band energy diagram, 451
- electroluminescence, 447–9
- growth, 446–7
- n-type, 449
- optical waveguides, 449–51
- quantum staircase emitter, 450
- Si/SiGe QC structure valence band, 448
- thick Si/SiGe QC structure top and bottom part, 447
- TM and TE polarised EL, 448

resistance ratio vs electron density, 377
- L = 50 Å for IRS (Δ = 3 Å, Λ = 30 Å), 377
- L = 100 Å with IS for different positions, 377

scattering time ratio vs electron density, 375
- L = 50 Å with IRS with Δ = 3 Å, 376

SiGe/Si quantum dot superlattice solar cell, 352–4
- 40-period QDSL, 354

strain engineering, 247–82, 403–7
- conduction and valence bands on strained Ge, 406
- conduction and valence bands on strained Si, 406
- growth insights, 250–60
- island engineering, 260–8
- potential applications, 272–9
- related information, 279–82
- rolled-up nanotechnology, 268–72

Stranski–Krastanov growth, 342

structural properties, 3–23
- critical thickness, 8–17
- crystal structure, 4
- future trends, 23
- lattice parameters, 4–7
- phase diagram, 7–8
- structural characterisation by X-ray diffraction, 17–23

surface orientation on defects, 343
  - dislocation showing stacking fault, 343
- Thompson tetrahedron application to extended defects, 346–9
  - biaxial compressive stress state conversion, 347
  - biaxial tension, 348
  - shear strain on a {111} glide plane, 348
- transport in quantum wells, 372–9
  - critical quantum well width L_c vs electron density, 379
  - lowest-order results, 372–7
  - metal–insulator transition, 378–9
  - mobility vs quantum well width, 378
  - resistance ratio vs electron density, 377
- transport properties, 361–93
  - comparison with experimental results, 383–9
  - future trends, 389–91
  - model, disorder and transport theory, 363–72
  - transport in heterostructures, 379–83
- transport properties and applications in devices, 399–428
  - future trends, 427–8
- transport properties of strained heterostructures
  - band and transport parameters, 402

silicon–germanium virtual substrates
- strain engineering, 147–67
  - compositionally graded buffer, 148–58
  - ion-implantation buffer, 161–5
  - low-temperature buffer, 158–61
  - other methods and future trends, 165–7

silicon hydrides, 53

silicon nitride, 554

silicon-on-insulator, 486
- demand for virtual substrate and (Si)Ge on insulator, 171–2
- dislocation formation and reduction mechanism, 180–5
  - conventional vs two step method, 184

- evaluation technique for dislocations in SGOI, 180–1
  - SGOI layers TEM images, 182
  - threading dislocations incidence during Ge condensation, 181–3
  - two-step Ge condensation method proposal, 183–5
  - whole SGOI layers plain-view images, 181
- extension toward Ge on insulator, 186–7
  - GOI layer TEM images, 187
  - Raman spectroscopy variation during Ge condensation process, 187
- Ge condensation method fundamentals, 172–5
  - Ge atoms depth profiles variation, 174
  - mechanism verification, 173–5
  - SIMOX technology limit, 172–3
- Ge profiles
  - illustration, 173
  - temperature dependence, 175
- growing silicon–germanium layers on thin substrate, 171–88
- SGOI formation by Ge condensation method, 172–86
  - fabrication method variation and strained-Si on insulator, 185–6
- strain relaxation during Ge condensation, 176–80
  - continuous SGOI layer having no boundary, 176–7
  - relation between relaxation ratio at the mesa centre and radius, 179
  - relation between relaxation ratio of SGOI structure, 177
  - SGOI layer shaped as a mesa island, 178–80
  - strain distribution along radial direction for SGOI mesa, 178

silicon oxynitride, 554

silicon(001) thin films
- germanium islands evolution modelling, 211–43
  - beyond the SK onset: SiGe intermixing, 232–5
  - beyond the SK onset: vertical and horizontal ordering for applications, 235–41
  - epitaxial growth modelling consideration, 211–13
  - future trends: ordering Ge islands on pit-patterned S(001), 241–3
  - Stranski–Krastanow heteroepitaxy onset, 219–32

SIMOX process, 172, 173, 176

SIMOX technology, 171, 172

single complete layer formation, 87

Si/SiGe, 477–9

smart cut, 185

solar cells, 339, 563–5
- plasmon solar cell structure, 564

solid-phase crystallisation, 192

solid-source method, 84–5

SPC *see* solid-phase crystallisation

spin-polarised electron gas, 367–8

spin-splitting, 413

spintronics
- nanostructured manganese dilute magnetic semiconductor, 575–602
  - electric field-controlled ferromagnetism, 596–601
  - ferromagnetism theories, 578–81
  - future trends, 601–2
  - growth and characterisations, 581–96

split-off band, 28

spreading resistance profiler, 317

SQUID *see* superconducting quantum interference device

STM *see* scanning tunneling microscopy

strain, 314–16
- compressive vs tensile, 345
- silicon on insulator, 185

strain engineering, 247, 403–7
- compositionally graded buffer, 148–58
  - crosshatch morphology, 151–5
  - Raman mapping images for strained-Si/SiGe virtual substrate structure, 156
  - relaxation mechanism, 149–51
  - strained Ge channel layers TEM images, 156
  - strain field fluctuation, 155–9

    growth insights, 250–60
        heteroepitaxy, 250–2
        hut clusters, 255
        island evolution, 252–60
        pyramid and a hut cluster, 253
        SiGe islands obtained after deposition, 256
        superdome, 257
    ion-implantation buffer, 161–5
        SiGe buffer layer image on Si-ion-implanted Si substrate, 164
    island engineering, 260–8
        laterally ordered SiGe/Si(001) islands, 266
        lateral ordering, 263–8
        vertical alignment of Ge dots, 260–3
        vertically aligned Ge islands, 261
    low-temperature buffer, 158–61
    other methods and future trends, 165–7
        Raman mapping image for SiGe buffer layer, 166
    potential applications, 272–9
        electronic transport and spin properties, 272–5
        freestanding microtubes and contact pads, 274
        thermal conductivities, 278
        thermal transport, 275–9
    rolled-up nanotechnology, 268–72
        roll-up process to create nanotubes, 269
        structure with individual wrinkling channels, 272
    SiGe virtual substrates
        strain distribution wavelength, 157
        surface AFM images, 154
        various virtual substrates, 149
    silicon–germanium micro- and nanostructures, 247–82
        Ge nanostructures on Si(001), 248
        micro- or nanotube schematic, 248
        potential applications, 272–9
        rolled-up nanotechnology, 268–72
    silicon–germanium virtual substrates, 147–65
strain feld, 155–9
Stranski–Krastanov growth mode, 15, 46, 50, 56, 87–8, 251, 254, 342, 421
Stranski–Krastanow heteroepitaxy, 213–35
    Ge islands STM image, 214
    Ge/Si(001) reconstruction, 218
    onset, 219–32
        AFM images sequence, 228
        3D islands energy difference and volume, 221
        Ge/Si(001)surface energy filled circles, 223
        Ge WL made of *N* Ge layers, 220
        metastable domes and pyramids, 226
        model limitations, 230–2
        PL spectra recorded at different coverages, 229
        RS-reconstructed view, 223
        simple thermodynamic model for SK growth, 219–21
        Si substrate elastic-energy distribution, 225
        surface and volumetric terms in Ge/Si(001), 221–5
        unexpected theoretical predictions *vs*. experiments, 225–30
    SiGe intermixing, 232–5
        Ge concentration profile evolution, 234
    strain maps for shallow Ge pyramid by 3D FEM calculations, 215
    Tersoff, Teichert and Lagally model, 235–41
        background, 235–7
        island distribution AFM image, 240
        islands layers ordering, 238
        islands region schematic process, 241
        islands vertical ordering, 237
        ordering from the first layer induced by moderate capping, 239–41
        shape effects, 237–9
        single and upper layer island AFM image, 236
        Si SL strained distribution and chemical potential, 239
substrate, 250
superconducting quantum interference device, 592
superdomes, 214
superlattices, 249

Super Photon ring 8 GeV, 95
supersaturation, 47
surface free energy, 48, 49

T-∑9, 88–90, 92–3
TDD *see* threading dislocation density
TED *see* transmission electron diffraction
Tersoff, Teichert and Lagally model, 235–41
  background, 235–7
  island distribution AFM image, 240
  islands layers ordering, 238
  islands vertical ordering, 237
  ordering from the first layer induced by moderate capping, 239–41
  shape effects, 237–9
  single and upper layer island AFM image, 236
  Si SL strained distribution and chemical potential, 239
TFTs *see* thin-film transistors
thermal transport, 275–9
thermodynamics
  crystal growth, 46–50
    Gibbs free energy, 49
thermoelectric figure of merit, 275
thin-film transistors, 191
Thompson tetrahedron, 10, 346–9, 355
threading dislocation density, 344
threading dislocations, 180, 182
Tiller's criterion, 74
$TiSi_2$ *see* titanium disilicide
titanium disilicide, 459–60
transceiver IC, 491
transient enhanced diffusion, 305, 320–2
transmission electron diffraction, 88
transmission electron microscopy, 17, 88, 180, 581
transmitter chipset, 491
travelling liquidus-zone method, 78
tri-gate MOSFETs, 517–20
two-dimensional electron gases, 407–19
  low-temperature mobility and metal-to-insulator transition, 407–10
  Quantum hall effect, 410–15
  Quantum point contacts, 415–17
  room-temperature electron mobility, 417–19
two-dimensional hole gases, 419–23
  pseudomorphic Ge quantum wells on Ge-rich metamorphic substrates, 421–3
  pseudomorphic SiGe quantum wells on unstrained Si, 419–21

ultra high vacuum chemical vapour deposition, 118
ultra-large-scale integrated (ULSI) applications
  electrical properties, 462–6
    contact resistivities of NiSi, $NiSi_2$, $CoSi_2$, and C54-$TiSi_2$, 464
    contact resistivities of p-type contact, 463
    Pd, Pt, Ir, and Er silicides, 465
  silicide and germanide technology, 456–66
    electrical properties, 462–6
    silicide and germanosilicide thin films formation, 457–8
    silicides crystalline properties, 459–62
  silicide and germanosilicide thin films formation, 457–8
    salicide formation in MOSFETs, 457
  silicides crystalline properties, 459–62
    $CoSi_2$, 460
    crystalline and electrical properties, 459
    NiSi, 460–2
    plan-view TEM of NiSi layer on Si, 461
    $TiSi_2$, 459–60
ultrametastable, 12
ultra-thin body, 501
ultra-thin oxide, 546
undercooling, 48
uniaxial compressive strain, 506
uniaxial strain, 505
uniaxial tension, 347

valence-band minima, 28
van der Merwe's group, 12
Vegard's law, 5, 12, 28, 104
virtual substrate, 147
Volmer–Weber modality, 213

Volmer–Weber mode, 15, 46, 50, 56, 251

wafer, 119, 121
Walton relation, 60
wavelength-division-multiplexing data transmission, 552
wetting layer, 213
WGMs, 438, 439, 445
Wiedemann–Franz law, 275
Wigner–Seitz parameter, 370
wireless communications, 491–3
  chip photomicrograph of single chip 5.8GHz, 493
  transmitter and receiver built with the IC chipset, 492

X'Pert Epitaxy, 19
X-ray diffraction, 6
  structural characterisation, 17–23
    different reflections plot recorded from strain-relieved SiGe layer, 23
    sample positions in X-ray beam, 20
    SiGe layer on Si(100) vs computer simulated SiGe on Si structure, 19
    single crystal heterostructure, 18
    strained layer inclined net planes and substrate, 21
X-ray micromodification, 94
X-ray reciprocal space mapping, 79
X-ray rocking curve, 17
X-ray topography, 17
XRMD *see* X-ray micromodification

Zener–kinetic exchange, 578
zero-field cooled magnetisation, 583, 592